The dynamics of transcript abundance during cellularization of developing barley endosperm

Zhang, Runxuan; Tucker, Matthew R.; Burton, Rachel A.; Shirley, Neil J.; Little, Alan; Morris, Jenny

Published in:
Plant Physiology

DOI:
10.1104/pp.15.01690

Publication date:
2016

The Dynamics of Transcript Abundance during Cellularization of Developing Barley Endosperm[1][OPEN]

Runxuan Zhang[2], Matthew R. Tucker[2], Rachel A Burton, Neil J. Shirley, Alan Little, Jenny Morris, Linda Milne, Kelly Houston, Pete E. Hedley, Robbie Waugh, and Geoffrey B. Fincher*

The James Hutton Institute, Invergowrie, Dundee, DD2 5DA, United Kingdom (R.Z., L.M., K.H., P.E.H., R.W.); Australian Research Council Centre of Excellence in Plant Cell Walls, School of Agriculture, Food and Wine, University of Adelaide, Waite Campus, Glen Osmond, SA 5064, Australia (M.R.T., R.A.B., N.J.S., A.L., G.B.F.); and Division of Plant Sciences, College of Life Sciences, University of Dundee, Dundee, DD1 4HN, United Kingdom (R.W.)

ORCID IDs: 0000-0003-4661-6700 (M.R.T.); 0000-0003-2136-245X (K.H.); 0000-0003-0866-324X (P.E.H.); 0000-0003-1045-3065 (R.W.); 0000-0003-4703-3494 (G.B.F.).

Within the cereal grain, the endosperm and its nutrient reserves are critical for successful germination and in the context of grain utilization. The identification of molecular determinants of early endosperm development, particularly regulators of cell division and cell wall deposition, would help predict end-use properties such as yield, quality, and nutritional value. Custom microarray data have been generated using RNA isolated from developing barley grain endosperm 3 d to 8 d after pollination (DAP). Comparisons of transcript abundance over time revealed 47 gene expression modules that can be clustered into 10 broad groups. Superimposing these modules upon cytological data allowed patterns of transcript abundance to be linked with key stages of early grain development. Here, attention was focused on how the datasets could be mined to explore and define the processes of cell wall biosynthesis, remodeling, and degradation. Using a combination of spatial molecular network and gene ontology enrichment analyses, it is shown that genes involved in cell wall metabolism are found in multiple modules, but cluster into two main groups that exhibit peak expression at 3 DAP to 4 DAP and 5 DAP to 8 DAP. The presence of transcription factor genes in these modules allowed candidate genes for the control of wall metabolism during early barley grain development to be identified. The data are publicly available through a dedicated web interface (https://ics.hutton.ac.uk/barseed/), where they can be used to interrogate co- and differential expression for any other genes, groups of genes, or transcription factors expressed during early endosperm development.

Cereal grains from the Poaceae family of higher plants are consumed by many societies around the world and in many cases provide a high proportion of the daily caloric intake required in human diets. The grain consists of an embryo, which develops into the new seedling following grain germination, and an endosperm, which contains storage carbohydrate, protein, and other nutrients that support the growth of the young seedling after grain germination, until it becomes photosynthetic and self-sufficient. A tough outer pericarp-testa layer surrounds and protects the embryo and endosperm from abiotic, biotic, and mechanical stresses. The endosperm consists of an inner starchy endosperm, which generally accounts for the largest proportion of the mature grain, and an aleurone layer that is one to a few cells in thickness and provides the battery of enzymes required for starchy endosperm mobilization following germination.

The endosperm and its nutrient reserves are particularly important factors both in biological terms and in the context of grain utilization. The total amount of starch, storage proteins, and to a lesser extent lipids and minerals, determine the nutritional value and quality of the grain, both for the human diet and in grain processing. These factors, together with the speed with which the reserves are released to support seedling growth after germination, are also important determinants of seedling vigor and hence crop establishment. The total number of cells in the endosperm is believed to influence grain size (Trafford et al., 2013).

There has been considerable interest in the genetics, physiology, and biochemistry of grain development, and that of the endosperm in particular. The endosperm is initiated through a double fertilization event, where a

[1] We acknowledge funding from the BBSRC (Response Mode grant BB/J019593/1) and the Scottish Government strategic program Work Package 5. This work was also supported by grants from the Australian Research Council. M.R.T. was supported by an ARC Future Fellowship.

[2] These authors contributed equally to this work.

* Address correspondence to geoff.fincher@adelaide.edu.au.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Geoffrey B. Fincher (geoff.fincher@adelaide.edu.au).

R.Z. performed the initial coexpression analyses and created the networks from the microarray data; M.R.T., N.J.S., K.H., and A.L. analyzed the modules and validated transcript data; P.E.H. and J.M. designed and processed the microarrays, and L.M. constructed the database; R.A.B. isolated the RNA from developing barley grain; R.W. and G.B.F. conceived the project and helped design the experiments; and M.R.T., R.Z., G.B.F., and R.W. wrote the manuscript.

[OPEN] Articles can be viewed without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.15.01690

sperm nucleus undergoes syngamy with two haploid polar nuclei in the central cell. The fertilized, triploid endosperm nucleus subsequently undergoes numerous rounds of free nuclear division in the highly vacuolated central cell, or syncytium, without cytokinesis (Paulson, 1969; Brown et al., 1994; Olsen, 2004; Wilson et al., 2006). The multiple nuclei are positioned around the periphery of the syncytium until about three days after pollination (DAP), when nascent cell walls start to grow around individual nuclei. This cellularization process occurs without mitosis, involves a primitive cytoplasmic phragmoplast, and initially leads to the formation of anticlinal cell walls, which extend centripetally from the perimeter toward the center of the developing caryopsis. The partitioned nuclei subsequently undergo typical mitotic divisions that form cell plates between the daughter nuclei and result in the deposition of periclinal cell walls. The alternating cycle of anticlinal cell wall growth and mitotic divisions leading to the deposition of periclinal cell walls is repeated several times until the entire volume of the endosperm becomes cellular, which is usually between four and six DAP (Brown et al., 1994; Olsen, 2004; Wilson et al., 2006). During this cellularization process, the nascent walls are comprised predominantly of the (1,3)-β-glucan, callose (Brown et al., 1994; Wilson et al., 2006). As the grain fills and expands, normal cell division continues and further differentiation occurs. In developing barley grain, aleurone cells are easily recognizable at 10 DAP and a histologically distinct subaleurone layer of starchy endosperm cells becomes evident by 14 DAP (Becraft and Asuncion-Crabb, 2000; Wilson et al., 2012). The aleurone cells remain triploid and living in the mature grain, but the starchy endosperm cells become polyploid through endoreduplication and finally undergo programmed cell death and are nonliving in the mature grain (Young and Gallie, 2000). The callosic material that was abundant in the early cell walls is replaced by arabinoxylans and (1,3;1,4)-β-glucans, which are the major noncellulosic wall polysaccharides in mature grains (Fincher, 1975; Bacic and Stone, 1981).

Endosperm development in grain of the Poaceae is therefore a complex process and involves cell division, cellular differentiation events, and the deposition of a wide range of storage polymers and other nutrients that change dramatically with time. In a previous study Sreenivasulu et al., (2008) examined the transition of transcriptional profiles of barley genes between grain maturation and germination and, in particular, examined the genes involved in hormonal responses and the mobilization of endosperm reserves after germination. Thiel et al. (2012) used laser-assisted microdissection (LAM) to identify mRNAs that are transcribed in barley endosperm transfer cells 3, 5, and 7 d after fertilization (DAF), and Pellny et al. (2012) have used RNAseq to examine genes involved in cell wall biosynthesis in developing wheat endosperm, but they examined transcript abundance after 10 DAP. Similarly, Pfeifer et al. (2014) examined global transcriptional profiles in wheat endosperm at 10 DAP, 20 DAP, and 30 DAP. In contrast, we have focused on the rapidly changing temporal events that occur very early in endosperm development from 3 DAP to 8 DAP, when both cell division and cell wall deposition occur in an unusual manner during syncytium cellularization, and when there is a dramatic change in the composition of the walls. In addition, we wished to define regulatory genes that might influence these unique cellular and biochemical processes through the identification of transcription factor (TF) genes that are cotranscribed with genes known to mediate cell-wall polysaccharide synthesis. The availability of information on the temporal and spatial appearance of wall polysaccharides during barley endosperm development (Wilson et al., 2006, 2012) has allowed the transcript analyses to be linked to morphological data and hence to other key events in grain development.

Thus, we have used sequential transcript profiling during early endosperm development to provide information not only on the genes involved in the diversity of genetic, biochemical, and cellular events that take place during these stages of grain development, including syncytium cellularization, initial cell division, stages of aleurone differentiation, starch synthesis and storage protein deposition, but also on genes that mediate the regulation of the processes. In this study, transcript analyses have been performed using a custom 44 K barley microarray to monitor the dynamic changes in mRNA abundance in barley endosperm sampled at daily intervals in the 3 DAP to 8 DAP developmental window. Molecular network analyses were then used to identify coexpressed gene clusters that are potentially under coordinated control. Although we have used the transcript abundance data to investigate specific temporal changes in genes involved in cell wall synthesis, remodeling, and degradation, together with the identification of transcription factors that might influence wall metabolism, our data are made publicly available as a resource for the analysis of other important processes during early endosperm development in barley.

RESULTS

Early Postfertilization Endosperm Development in Barley

Developing barley grains were collected from 3 DAP to 8 DAP and examined by transverse sectioning, toluidine blue staining, and light microscopy to determine the temporal sequence of developmental events. In conjunction with results from previous studies (Olsen, 2004; Wilson et al., 2006; Sreenivasulu et al., 2010), these were used to create a schematic that documents the early stages of barley grain development (Fig. 1). Under the growth conditions used in this study, the syncytial phase of endosperm development continued until approximately 3 DAP, at which point the nuclei were evenly arranged at the periphery of the embryo sac. Endosperm cellularization began as early

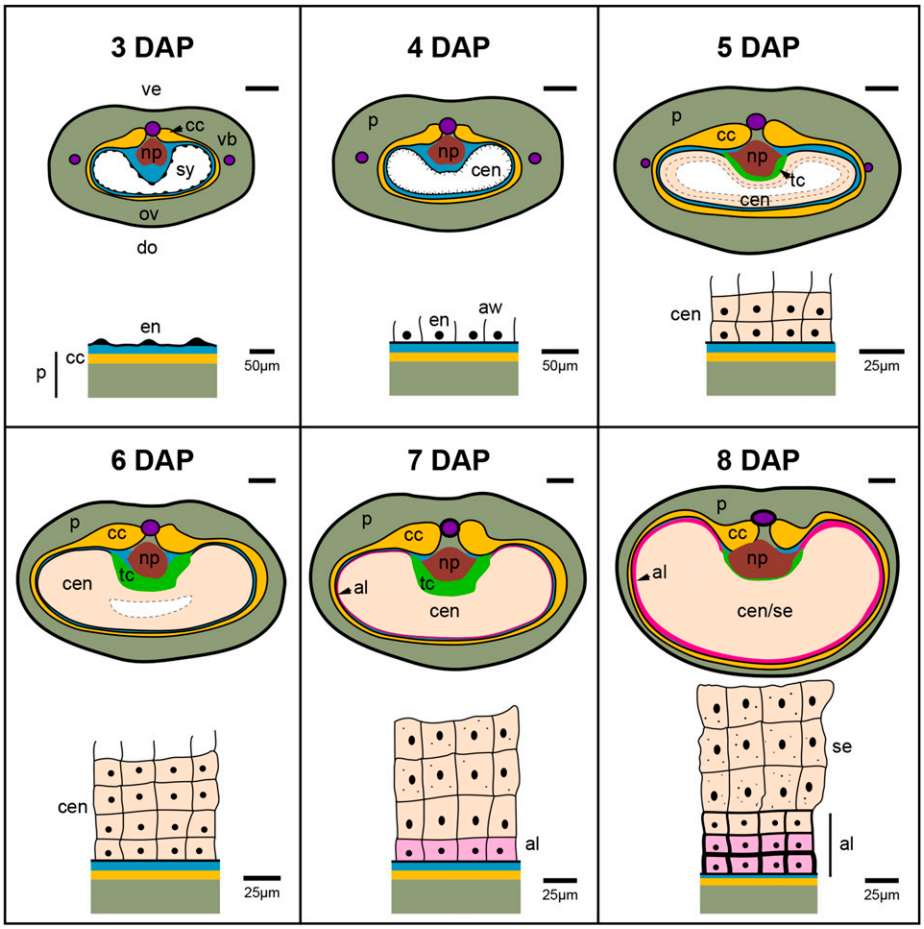

Figure 1. Diagram summarizing the developmental stages of barley grain and endosperm during the period 3 DAP to 8 DAP in barley (cv Sloop). Schematic representations of transverse sections from the midpoint of the grain are shown as well as a magnified view of the early events of endosperm development. At 3 DAP, the pericarp makes up the major part of the ovule, but the syncytium is evident and unenclosed nuclei can be detected around the periphery of the embryo sac. At 4 DAP the syncytium is still expanding and the first outgrowths of the anticlinal walls can be seen. By 5 DAP anticlinal walls have formed, the first enclosed nuclei can be seen around the periphery of the endosperm, and cellularization is almost complete. The cellularized endosperm continues to expand at the expense of the pericarp and differentiation of a single layer of aleurone cells has occurred by 7 DAP. Starch granules are also visible from 7 DAP onwards. At 8 DAP the starchy endosperm and the surrounding immature aleurone layer, which is two to three cells in thickness, are present. Bar = 250 μm unless otherwise indicated. (al, Aleurone; aw, anticlinal walls; cc, chlorenchyma layer; cen, cellular endosperm; do, dorsal; en, endosperm nuclei; np, nucellar projection; ov, ovary; P, pericarp; se, starchy endosperm; sy, syncytium; tc, transfer cells; vb, vascular bundle; ve, ventral.)

as 4 DAP when cell walls initiated between sister and nonsister nuclei, perpendicular (i.e. anticlinal) to the syncytial cell wall. From 5 DAP to 7 DAP, additional rounds of cell division were observed, starting with the formation of periclinal walls approximately parallel to the outer syncytium and followed by mitotic divisions and eventually a completely filled embryo sac. From 6 DAP to 7 DAP onwards, radial symmetry became apparent with the differentiation of the aleurone progenitor cells from cells in the outer starchy endosperm. The characteristic morphology of aleurone cells was evident at 8 DAP. From 7 DAP onwards starch granules could be detected in the inner cells of the cellularized endosperm. This developmental time course aligned closely with that of a previous study on cell wall polysaccharide accumulation during barley grain development (Wilson et al., 2006).

Transcriptional Analysis of Early Endosperm Development in Barley

Endosperm samples were harvested by manually dissecting the contents of the developing barley embryo sac and peripheral cytoplasm from 3 DAP to 8 DAP (Supplemental Fig. S1). Total RNA was isolated from three biological replicates (n = >50 grains for each stage) and analyzed using a custom Agilent 44 K microarray platform (Agilent Technologies, Santa Clara, CA). A primary analysis of the mRNA expression

profiles for these samples using principal component analysis (PCA) showed gradual changes in transcript expression profiles over time (Supplemental Fig. S2). The total number of probes recording gene expression in each sample varied from 19,424 to 25,558 (Table I) and these are hereafter referred to as "expressed probes". To capture temporal changes, we compared adjacent time points (for example, 4 DAP versus 3 DAP and 5 DAP versus 4 DAP) using a Welch's *t* test (p-value $\leq$ 0.01) on an ANOVA filtered list, using an FDR of 1×10^{-4} (Supplemental Table S1), of probes showing differential expression between any two time points. This analysis generated lists of differentially expressed genes that varied in number from 32 (8 DAP versus 7 DAP) to 785 (5 DAP versus 4 DAP; Table I). The largest change in differential gene expression occurred between 4 DAP and 5 DAP and was consistent with the original PCA plot (Supplemental Fig. S2).

We subsequently explored changes in the abundance of transcripts from genes that mediate known molecular functions or cellular processes at the different stages of development, using gene ontology (GO) enrichment analysis. Examples of what we refer to as "molecular functions" include cellulose synthesis, starch synthesis, cell cycling, and regulation by transcription factors. The microarray probe sequences were compared to barley Morex locus (MLOC) transcripts (predicted from the barley reference genome assembly; Mayer et al., 2012) via BLAST searches with an e-value cutoff of 1e-20 and the GO annotations of each probe were assigned using the GO terms of associated barley MLOCs, which had been determined in a previous study (Baker et al., 2014). All probes on the microarray were investigated for enriched GO terms using the R Bioconductor package TopGO (TopGO v. 2.12, R v. 3.1.0) and the Fisher test with weight algorithm (p-value $\leq$ 0.05). The values shown in the GO enrichment tables (e.g. Supplemental Table S2) are enrichment p-values of the GO term in the corresponding gene list.

For the differentially expressed genes (Supplemental Table S1 and Supplemental Table S2), GO analysis identified molecular functions that were significantly enriched at each time point. For example, the 4 DAP versus 3 DAP gene list revealed enrichment for terms associated with cell wall remodeling, such as glycosyl hydrolase and pectin methylesterase (GO:0016798 and GO:0030599). Enriched terms in the 5 DAP versus 4 DAP list included genes associated with the metabolism of proteins, such as peptidase inhibitors (e.g. GO:0004867, GO:0004866, GO:0061135, GO:0030414, and GO:0061134), and terms linked to cell wall biosynthesis, such as glycosyl transferase (GO:0016759 and GO:0016757). Similarly, the 6 DAP versus 5 DAP set was enriched in terms that are linked to peptidase activity and to starch biosynthesis (for example, GO:0003844). Molecular functions enriched in the 7 DAP versus 6 DAP set included transmembrane transport (GO:0003924, GO:0005342, GO:0046943, and GO:0008514), while in the 8 DAP versus 7 DAP group, regulators of peptidase and enzyme activity were once again abundant (GO:0030414, GO:0061134, and GO:0004857).

Probes that were preferentially expressed, which we defined as being expressed in at least two out of three biological replicates exclusively at one time point but not at any other time points, were also investigated (Table I, Supplemental Table S3). The number of preferentially expressed probes ranged from 614 at 4 DAP to 8 at 7 DAP, with the majority detected at 4 DAP, 5 DAP, and 6 DAP. The 5 DAP sample showed the greatest abundance of GO terms (Supplemental Table S4). Of particular note at 5 DAP were molecular function terms that included transmembrane receptor signaling (e.g. GO:0004888), hormone transport (e.g. auxin efflux, GO:0010329) and cellulose biosynthesis (e.g. GO:0016759).

Molecular Network Analysis of Genes Showing Differential Transcript Accumulation

In total, 5925 probes showed significantly different expression between any two time points in the six stages of endosperm development, using an FDR threshold of 1×10^{-4} (Supplemental Table S1). The expression profiles of these probes can be accessed through the Web site (https://ics.hutton.ac.uk/barseed/). A gene coexpression network was constructed using the normalized microarray expression data of these 5925 probes from all 18 samples (three biological replicates for six time points) using a weighted gene coexpression analysis (Langfelder and Horvath, 2008) R library. Using a set of parameters that produce refined clusters, we detected 47 modules containing between 36 and 686 probes each (Table II). All of the modules, defined using color codes, are shown in Supplemental Figure S3. GO enrichment analyses of the gene lists for each module were conducted and significantly enriched GO terms (p-value $\leq$ 0.05) were identified. These are

Table I. *Microarray analysis of barley endosperm tissues from 3 DAP to 8 DAP*

Category	3 DAP	4 DAP	5 DAP	6 DAP	7 DAP	8 DAP
Number of expressed probes	23,179	25,918	25,558	25,403	19,424	20,745
Number of *x* DAP versus *x-1* DAP differential probes*	—	107	785	241	211	32
Number of preferential probes	90	619	215	283	8	31

*T-test: p-value $\leq$ 0.01

Table II. Number, name, and composition of modules identified from transcript accumulation patterns during early endosperm development in barley

Module No.	Module Name	No. of Probes (CUSTs)	Module No.	Module Name	No. of Probes (CUSTs)	Module No.	Module Name	No. of Probes (CUSTs)
1	Bisque 4	39	17	Green	268	33	Purple	118
2	Black	214	18	Green yellow	116	34	Red	239
3	Blue	686	19	Gray 60	95	35	Royal blue	91
4	Brown	425	19	Ivory	45	36	Saddle brown	139
5	Brown 4	40	21	Light cyan	98	37	Salmon	103
6	Cyan	102	22	Light cyan 1	45	38	Sienna 3	62
7	Dark green	84	23	Light green	92	39	Sky blue	71
8	Dark gray	79	24	Light steel blue 1	88	40	Sky blue 3	60
9	Dark magenta	67	25	Light yellow	92	41	Tan	104
10	Dark olive green	135	26	Medium purple 3	54	42	Thistle 1	36
11	Dark orange	72	27	Midnight blue	100	43	Thistle 2	37
12	Dark orange 2	41	28	Orange	78	44	Turquoise	505
13	Dark red	85	29	Orange red 4	56	45	White	71
14	Dark slate blue	246	30	Pale turquoise	68	46	Yellow	394
15	Dark turquoise	81	31	Plum 1	57	47	Yellow green	61
16	Floral white	41	32	Plum 2	38	—	—	—

summarized for each module in Supplemental Table S5. We also calculated the similarity between modules, using correlations of their eigenvalues, to cluster modules into related-module groups. These relationships are shown as a dendrogram and heat map in Figure 2.

Functional Subnetworks Identified through GO Analysis

Each of the 47 modules containing differentially expressed genes was examined for enrichment of GO classes (Supplemental Table S5). The aim was to determine whether the different modules might reflect temporal expression trends, and eventually identify individual candidate genes, that correlate with the key events of early endosperm development in barley. This was clearly the case (Supplemental Table S5, Supplemental Fig. S3). Alignment of the GO terms with the temporal arrangement of the 47 modules identified 10 broad clades (Fig. 2B), which covered many terms from energy catabolism to cell division and accumulation of storage reserves. Specific terms relevant to early grain development, such as chromatin regulation (turquoise and blue modules; peaks at 3 DAP to 5 DAP), cytokinesis (cyan and red modules; 3 DAP and 4 DAP to 5 DAP, respectively), cell division (red and light yellow modules; 3 DAP to 4 DAP and 6 DAP, respectively), mitosis (light cyan module; peaks at 5 DAP to 6 DAP), and starch synthesis (saddle brown and purple modules; 6 DAP to 8 DAP) were identified and can be used to predict key regulatory genes involved in these processes (Supplemental Table S5, Supplemental Fig. S3).

Terms related to cell wall synthesis and remodeling were also abundant, particularly in the yellow, red, plum 2, light cyan 1, and midnight blue modules that show transcript peaks at 4 DAP, 5 DAP, 6 DAP, and 8 DAP, respectively (Supplemental Fig. S3). Although the events of cell wall formation and polysaccharide deposition during grain development have been thoroughly investigated by immunohistological analysis in barley, the molecular basis for and the regulation of these dynamic changes has remained elusive. Accordingly, we focused on the further analysis of modules and genes associated with cell wall formation in attempts to identify key stages, interactions, and regulators in this critical process during grain development.

Cell Wall-Related Subnetworks during Early Grain Development

In order to provide confidence that our tissue sampling and gene expression analyses reflected changes in endosperm development, we first examined our datasets for signals that may indicate contamination by other developing grain tissues. This revealed that in several modules showing increased transcript abundance at 3 DAP, GO terms associated with photosynthesis were enriched (Supplemental Fig. S3). Since photosynthetic activity in early developing cereal grains is restricted to the pericarp layers that surround the developing endosperm and embryo, this indicated that some maternal RNA may have been collected with the 3 DAP endosperm sample. To estimate how many of the differentially expressed genes were endosperm-derived, we compared our gene lists with 454 sequence

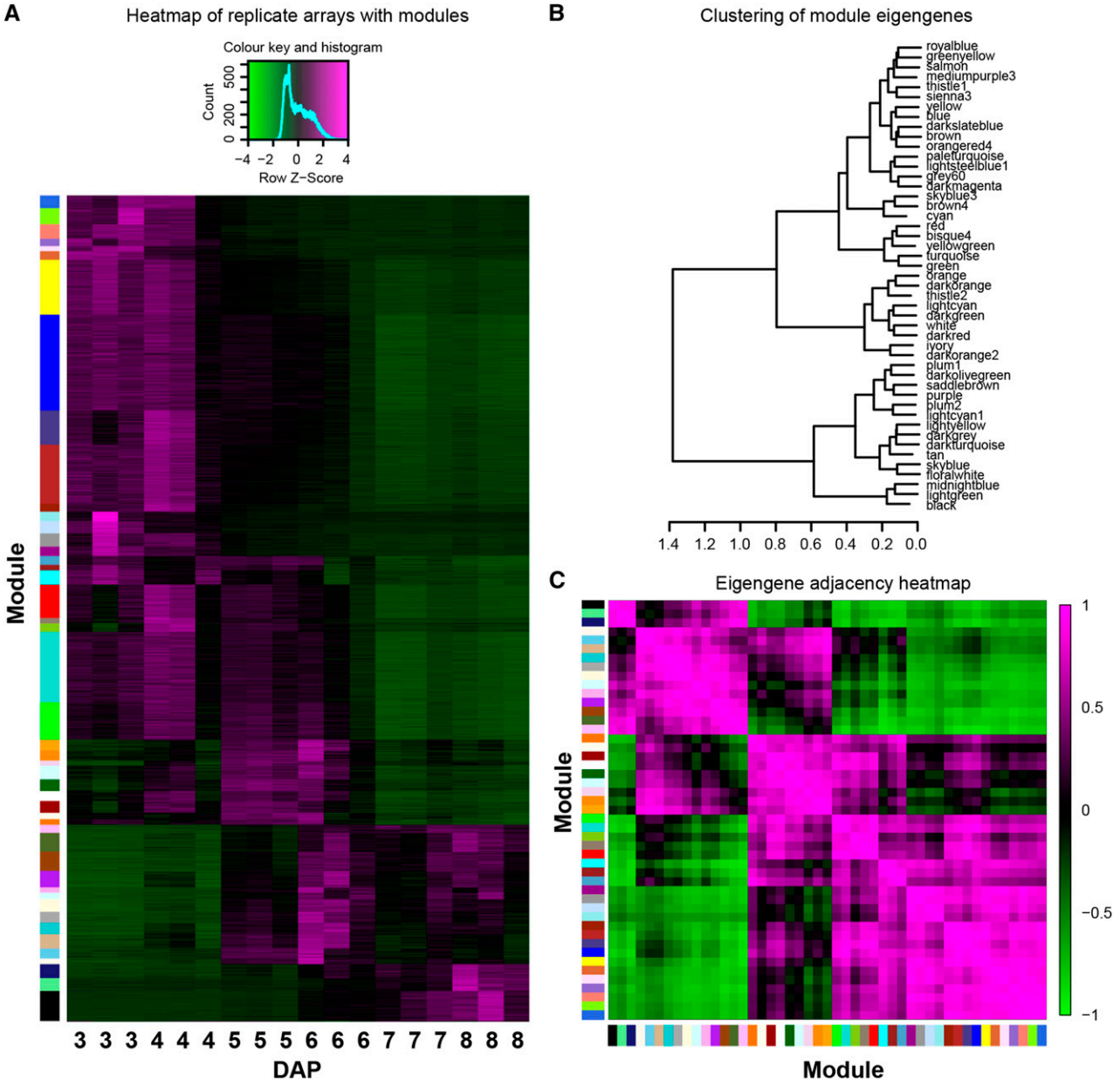

Figure 2. Transcript analysis during early barley endosperm development. A, Heatmap showing relative transcript abundance of each differential probe, arranged in a temporal sequence. Modules are indicated by color. B, Dendrogram identifying related clusters of modules, organized in the same order as (A). C, Correlation of different module eigenvalues, showing the similarity between each module.

data generated from LAM barley endosperm tissues collected at 3, 5, and 7 d after fertilization (DAP; Thiel et al., 2012). Although the entire endosperm was not harvested by LAM, the time points overlapped closely with several of the stages used in this study. The 454 data from Thiel et al. (2012) were mapped to barley MLOCs and custom microarray probes (CUSTs) in CLC Genomics, thereby allowing MLOC/CUST abundance to be determined within the LAM endosperm data. In total, 67% of the differential genes identified in this study could be aligned with LAM endosperm-specific 454 reads. Of the remainder, 23% of the total genes showed peaks in transcript abundance at 4 DAP or later by microarray analysis, when the presence of maternal tissue was less likely.

We also generated microarray data for whole barley caryopses, containing both pericarp and embryo sac/endosperm tissue, at 0 DAP, 1 DAP, and 2 DAP (Supplemental Table S6). Genes that were highly abundant at 3 DAP but less so at any other stage of caryopsis or endosperm development were considered as putative endosperm-expressed stage-specific genes

(approximately 3% of all differentially expressed genes). By contrast, genes that showed a peak in transcript abundance at 3 DAP in the endosperm dataset but were more abundant at 0 DAP, 1 DAP, and 2 DAP in samples that contained pericarp, were considered as possible maternal genes (~7% of all differentially expressed genes). Consistent with this hypothesis, a subset of these genes carried a GO annotation related to photosynthesis (Supplemental Fig. S6). However, in the vast majority of cases, the orthologous rice sequences for these putative maternal genes were detected in rice endosperm tissues at 3 DAF, 9 DAF, or 12 DAF (https://genevestigator.com/gv/doc/intro_plant.jsp; Supplemental Fig. S4), suggesting that if the expression patterns are conserved, they may also represent endosperm-expressed genes in barley. Taken together, these analyses indicate that at least 93% of the genes showing differential expression in the barley endosperm stages analyzed here are likely to represent endosperm-expressed genes that vary between different stages of endosperm development.

In a second approach to defining the level of purity of the RNA samples, transcription patterns of marker genes for nucellus cell layers (Doan et al., 1996; Sturaro et al., 1998; Linnestad et al., 1998) were monitored. For the *HvEX1* and *HvNUC1* nucellus marker genes, transcript levels decreased dramatically from 2 DAP to 3 DAP. At the first point of the analysis (3 DAP), transcript levels of both genes were approximately 5% or less of the peak levels seen at 2 DAP (Supplemental Fig. S8). This is consistent with the conclusion above that the level of contamination of the endosperm RNA preparation at 3 DAP with nonendosperm RNA was less than 7% and that this level was even lower in the RNA from later time points.

The dynamic transcript abundance patterns observed in the yellow, red, plum 2, and light cyan 1 modules indicated that genes involved in cell wall biogenesis were differentially expressed at multiple stages during the early endosperm time course (Supplemental Fig. S3). The yellow module showed an early peak of expression at 3 DAP to 4 DAP where, based on GO annotation, genes involved in (1,3)-β-glucan (callose) biosynthesis were enriched. Following this, the red module showed a peak of expression at 4 DAP to 5 DAP, and included GO terms such as cell wall organization, biogenesis, and assembly. Light cyan 1 overlapped somewhat with red in terms of cell-wall organization terms, but showed a peak of expression at 6 DAP. Plum 2 showed a similar peak of expression to light cyan 1 at 6 DAP, and included additional GO terms such as primary cell wall biosynthesis and cell wall thickening. Notably, terms associated with sugar-nucleotide, polysaccharide, and glucan biosynthesis, which might be expected to be associated with cell wall formation, were also enriched in the green (3 DAP to 5 DAP), turquoise (3 DAP to 5 DAP), saddle brown (6 DAP), and dark olive green (6 DAP to 8 DAP) modules.

Spatial Representation of Cell Wall-related Networks during Early Endosperm Development

The module analysis provided a clear delineation of gene expression patterns and potential functional units during early endosperm development. Despite this, there were obvious overlaps between modules in terms of the expression dynamics and enriched GO terms. For example, correlation analysis using eigenvalues for each module confirmed the close relationships between several modules enriched for GO terms related to cell wall formation. The red and turquoise modules overlapped during 3 DAP and 4 DAP and were highly correlated (0.96, Pearson correlation; Fig. 2C), while the light cyan 1 and plum 2 modules showed a correlation of 0.94, overlapping during 5 DAP to 6 DAP (Fig. 2C). The changes in gene expression in different modules were sometimes highly similar, so we considered that specific relationships between genes might also be interpreted through network analysis of all 5925 differential probes. This was achieved using a Spearman correlation matrix and Cytoscape (Shannon et al., 2003). The spatial representation of all the coexpression data resulted in a complex network that was ball-shaped and essentially uninterpretable, due to the vast number of nodes and connections, even at high stringency.

Cell Wall-related Subnetworks Are Arranged in a Temporal Sequence that Separates Specific Gene Activities

To overcome this, the entire differential probe list was filtered for genes involved in any aspect of cell wall biosynthesis. This list included predicted cell wall-related genes from the CAZy web site and the Cell Wall Navigator (Girke et al., 2004). Around 600 genes are annotated as having a cell wall-related function in the barley genome (Mayer et al., 2012), and of these, 304 match CUSTs on the 44 K Agilent chip. The probes were filtered using CUST to MLOC best-BLAST, and of the 227 CUSTs that matched a single MLOC, 89 were differentially expressed during early grain development and therefore appeared within a module (Supplemental Table S7). Eight of the nine modules that were previously identified as being enriched for cell wall-related GO terms were represented, and accounted for 29 genes. An additional 28 modules contained between one and six cell wall-related genes. The coexpression network analysis of these 89 probes/MLOCs was filtered to show only the top 5% of correlations (Fig. 3).

The spatial representation of the coexpression data revealed a network of cell wall genes that was arranged in a temporal sequence, connected by one or a few nodes. This temporal map could be divided into two main networks, namely early (Fig. 3) and late (Supplemental Fig. S4). During the early stages of endosperm development at 3 DAP to 4 DAP (Fig. 3, A, B, and C), a cluster of genes encoding expansins, arabinogalactan proteins (AGPs), and hydrolytic enzymes

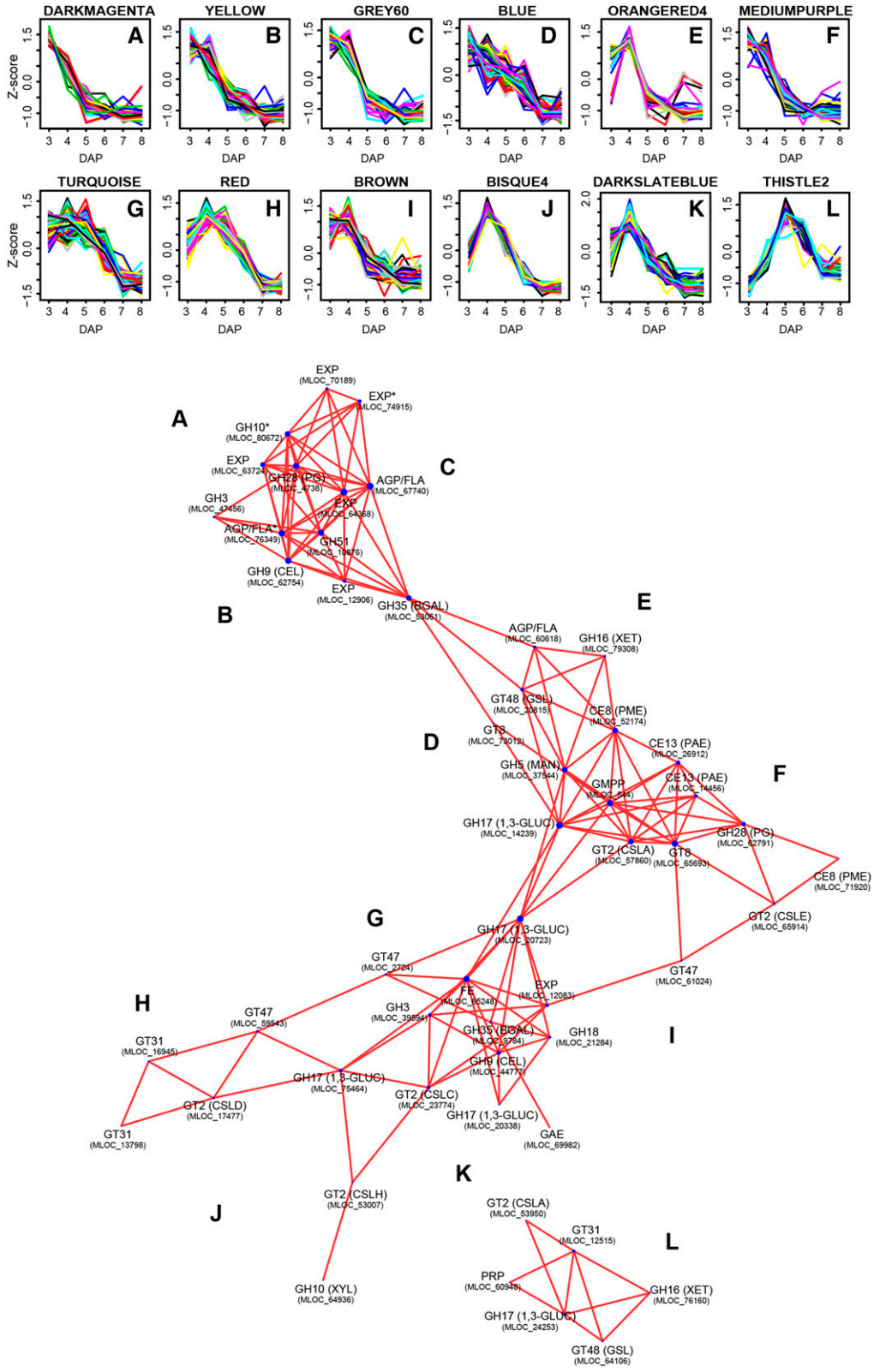

Figure 3. Spatial representation of coexpressed cell wall-related genes that are differentially expressed at 3 DAP to 5 DAP of early endosperm development. A, B, C, D, E, F, G, H, I, J, K, L, Different modules that are represented by genes in the network analysis. The relative position of these modules is indicated on the map. Asterisks indicate genes for which coexpression may be influenced by maternal transcript levels.

were coexpressed. This cluster was connected via a β-galactosidase to a second larger cluster, which contained genes that were most highly abundant around 3 DAP and 4 DAP (Fig. 3, D, E, and F), and contained a higher proportion of putative biosynthetic glycosyl transferase genes. A putative callose synthase (GT48), two glucuronyltransferases (GT8), a cellulose-synthase-like E (*CslE*; GT2), and a *CslA* (GT2) gene were identified. Several genes encoding hydrolytic- and polysaccharide-modifying enzymes, including (1,3)-β-glucanase (GH17), pectin methylesterase (PME), and pectin acetylesterase, were also present in this cluster. Further downstream, one of the (1,3)-β-glucanase (GH17) genes acted as a connecting node to a cluster (Fig. 3, G, H, I, J, and K) of modules enriched at 4 DAP and 5 DAP. This cluster showed another shift in gene identity, containing several putative xylosyltransferases (GT47), galactosyltransferases (GT31), and cellulose synthase-like genes (*CslD*, *CslC*, and *CslH*), together with genes encoding hydrolytic enzymes, such as a putative (1,3)-β-glucanase, a cellulase (GH9), and a xylanase (GH10). A stand-alone cluster (Fig. 3L) was also identified and was enriched for genes showing up-regulation at 5 DAP. This contained a diverse set of genes encoding putative enzymes implicated in callose biosynthesis and hydrolysis (GT48 and GH17), xyloglucan transglycosylation (XET), mannan biosynthesis (*CslA*), AGP biosynthesis (GT31), and cell wall structural proteins (Pro-Rich-Protein).

Genes that were differentially expressed in the subsequent stages of endosperm development (5 DAP onwards) clustered together in a less dense network that was also arranged in a temporal manner (Supplemental Fig. S4). At 5 DAP to 6 DAP (Supplemental Fig. S4A), the first GT34-encoding gene, which could be a xylosyltransferase or a galactosyltransferase, was identified and was associated with genes including a putative *CslA*, a GT47, an AGP, and two GH17 enzymes. At 6 DAP (Supplemental Fig. S4, B and C) a group of genes implicated in aspects of xylan biosynthesis was coexpressed. This group included a UDP-Xyl synthase and a UDP-Xyl epimerase, together with enzymes from the GT34, GT43, and GT61 (xylosyltransferase or arabinosyltransferase) families. Finally, at 6 DAP to 8 DAP (Supplemental Fig. S4, D, E, and F), genes encoding another GT47, a PME, a polygalacturonase (GH28), and an expansin were identified, as well as two starch biosynthetic genes (ADPglyphosS and ADPglyphosL) that act as markers for the deposition of starch during endosperm development.

Several of the close associations between cell wall-related genes identified in the network analysis appeared to be conserved in other barley tissues, as assessed by querying published RNAseq datasets (Fig. 4; Mayer et al., 2012). This suggested that some of the coexpressed genes may be regulated by similar transcriptional mechanisms across development, while other genes have been specifically recruited during early endosperm development to participate in the cellularization process.

In summary, the spatial representation of coexpressed cell wall-related genes revealed key associations between genes potentially involved in the biosynthesis and modification of endosperm cell wall polysaccharides. Some of the activities mediated by these genes have previously been linked to the biosynthesis of specific polysaccharides, but our analysis also identified candidate genes that have not previously been linked to endosperm cell wall synthesis.

TFs Associated with Temporal Events in Cell Wall Biosynthesis

The discovery of potential interactions between genes by network analysis led us to consider whether transcriptional regulators controlling different stages of cell wall biogenesis or modification might be identified in a similar manner. To test this, a list of putative TFs present within the barley genome was assembled and filtered in a similar manner to the differentially expressed cell wall genes described above. A total of 57 putative TF-encoding genes (Supplemental Table S8) were included, together with the 89 cell wall genes, to create a network map and the top 1% of correlations is shown in Figure 5. Four clusters were identified that were of particular interest. An AGAMOUS-LIKE TF (MADS-TF) that was coexpressed with a β-galactosidase, expansin, cellulase (GH9), and several additional TF genes was identified. These cell wall-related genes showed a similar transcript pattern that was enriched during the earliest stages of endosperm development analyzed (Fig. 5A). Another cluster (Fig. 5B) included a HD-ZIP TF, a *CslA*, a GH17, and a BFT3-like TF, all of which were most abundant at 3 DAP and 4 DAP of endosperm development. A third cluster (Fig. 5C) was abundant at 4 DAP, and included another MADS-TF as well as an AGP, a callose synthase (GT48), and a xyloglucan endotransglycosylase (GH16). Finally, a distinct cluster (Fig. 5D) containing a putative ethylene-response TF and five genes encoding putative sugar biosynthetic enzymes (GDH, UDP-Xyl synthase, UDP-Xyl epimerase, and GDP-Man pyrophosphorylase (GMPP)) was identified, and transcripts in this cluster peaked at 6 DAP.

DISCUSSION

In preparation for the transcript analyses during barley endosperm cellularization, it was important to ensure that the developmental stage of the grain, experimental procedures for RNA extractions, and the transcript profiling methods were clearly defined, reproducible, and consistent with previous studies. These criteria could be met through comparisons of the progress observed here for barley grain development with previously published studies in which identical conditions were used (Wilson et al., 2006, 2012). Moreover, it was possible to compare transcript profiles generated in this work by 44 K custom microarray analyses with

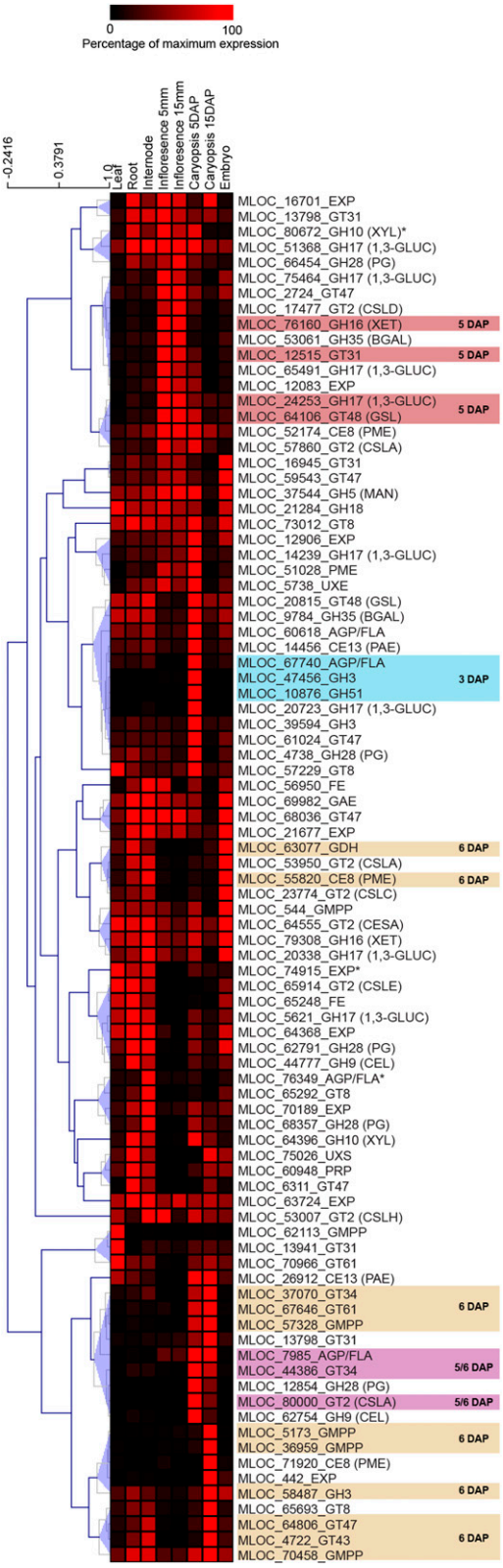

Figure 4. Analysis of cell wall-related transcripts during barley development. Transcript abundance was determined through analysis of RNAseq datasets from eight different tissue types. Values show the percentage of maximum transcript accumulation. Shaded boxes and DAP values indicate genes that were also found in the same or closely related modules during early endosperm development. Asterisks indicate genes for which coexpression may be influenced by maternal transcript levels.

data generated by quantitative RT-PCR, 22 K microarrays (Supplemental Fig. S5, A, B, C, D, E, and F′), LAM (Supplemental Table S9), and RNAseq (see Fig. 4; Burton et al., 2008; Sreenivasulu et al., 2008; Mayer et al., 2012; Wilson et al., 2012; Thiel et al., 2012). The comparisons showed that it was indeed possible to produce consistent transcript profiles with different methods and to accurately reproduce the conditions of grain development. Comparisons with barley endosperm LAM samples (Thiel et al., 2012) and with rice endosperm transcript databases (Nie et al., 2013; Sato et al., 2013) suggest that approximately 93% of the differentially expressed genes identified in this study are expressed in the endosperm. This is likely to be an underestimate, considering that the overlapping LAM data from Thiel et al. (2012) is almost certainly missing genes expressed in some central and peripheral regions of the developing endosperm. Moreover, some of the genes considered as possibly of pericarp origin at 3 DAP (~7% of the total) may be expressed in both maternal and endosperm tissue types, similar to orthologous genes from rice (Supplemental Fig. S6). Three expansin genes from this category were directly analyzed by qPCR, and two showed clear expression in barley endosperm tissues at 6 DAP, 8 DAP, and 12 DAP (Supplemental Fig. S7A). In total, 86 of the 89 (94%) cell wall genes and 51 of the 57 (90%) transcription factors identified as being differentially expressed and used for molecular network analysis could be independently verified as being endosperm-expressed. Therefore we are confident that good quality RNA can be extracted specifically from the endosperm with low levels of contamination with RNA from maternal tissues, such as the pericarp and nucellus, even very early after pollination (3 DAP) when the endosperm is still at the milky stage of development.

The availability of a tightly linked, daily time course of developing barley grain from 3 DAP to 8 DAP enabled detailed transcript analyses over a period that spanned many critical and progressive cellular and biochemical changes, including embryo sac enlargement and syncytium formation, anticlinal and periclinal wall formation, cellularization, and aleurone differentiation (Fig. 1). During this period, distinct cell-wall types appeared and disappeared and, toward the end of the period, starch and storage protein deposition had begun (Supplemental Fig. S1). PCA of the different time points each day from 3 DAP to 8 DAP showed clear separation of the data collected each day (Supplemental Fig. S2). By analyzing this data it was possible to sort the temporal variation of individual gene transcripts into 47 distinct groups, or modules, and the clear differences in the expression patterns of these are presented in Supplemental Fig. S2. Several modules show

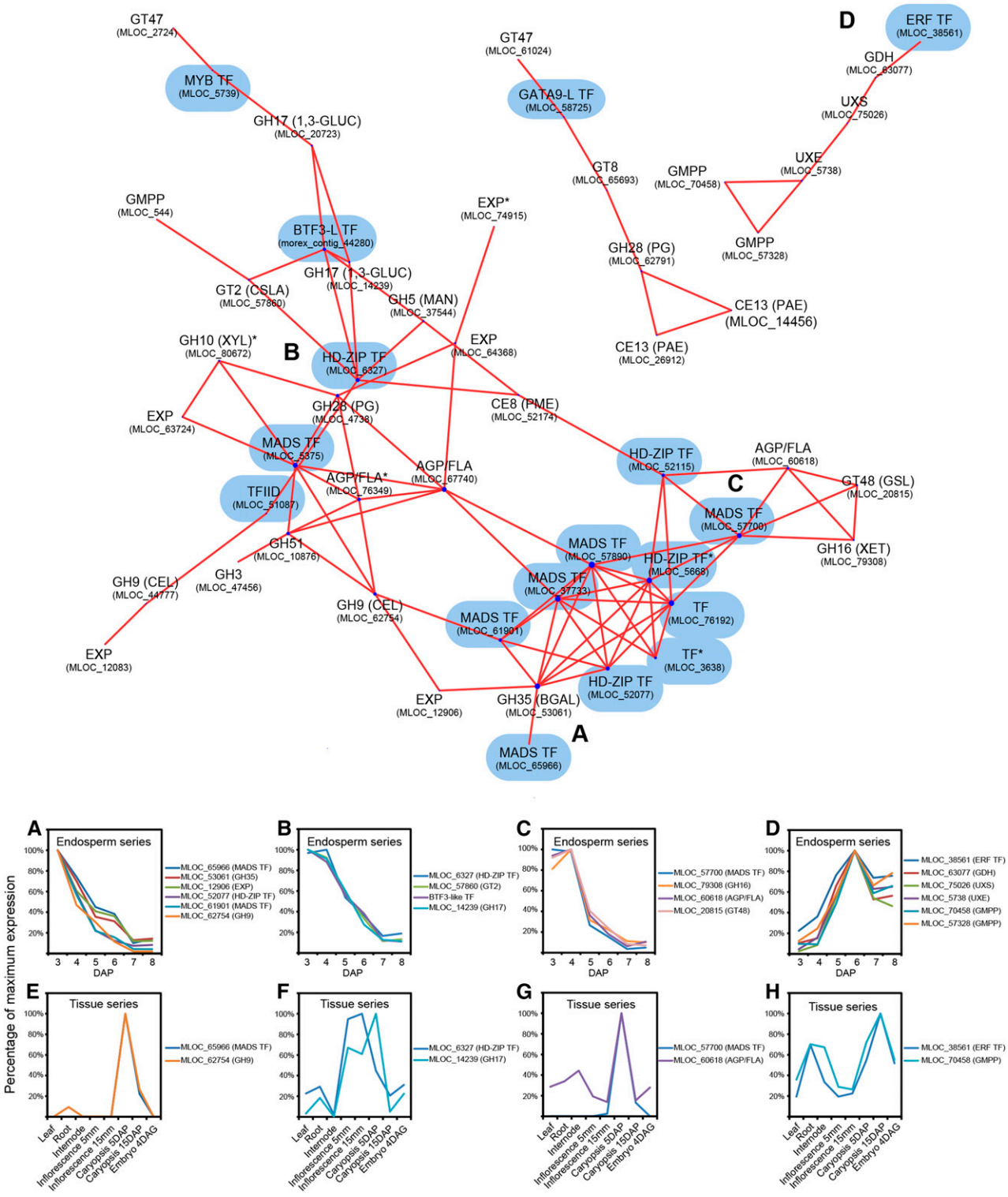

Figure 5. Spatial representation of coexpressed transcription factors and cell wall-related genes. A, B, C, D, Patterns of transcript abundance of selected genes during endosperm development based on custom microarray data. Values are shown relative to the maximum expression level. The spatial arrangement of the genes is indicated on the network map. E, F, G, H, Patterns of transcript accumulation for selected genes from (B–E) showing similar transcript accumulation patterns during general barley development based on RNAseq data. Values are shown relative to the maximum expression level. Blue shading on the network map highlights genes encoding TFs. Asterisks indicate genes for which coexpression may be influenced by maternal transcript levels.

patterns in which the level of differential expression decreases during the 3 DAP to 8 DAP period, others show a peak of expression around 4 DAP to 6 DAP, while still others show increasing levels of expression during the entire period of grain development examined (Supplemental Fig. S3).

Our next objective was to assess the possibility that individual genes or gene products within a module might be involved in specific molecular, cellular, or broader biological processes. Ontologies were superimposed on individual genes using GO enrichment (Ashburner et al., 2000; Baker et al., 2014) and, as expected, modules did show enrichment in GO terms related to processes such as energy catabolism, cytokinesis, starch synthesis, cell wall biosynthesis and degradation, transmembrane transport, and hormone metabolism. The GO enrichment analysis grouped the 47 modules into 10 related groups or clusters (Fig. 2) and the temporal similarities in the expression patterns for many of the modules in each cluster can be clearly seen (Supplemental Fig. S3). Some transcripts involved in photosynthesis were enriched in the 3 DAP modules (Supplemental Fig. S7B), indicating some level of maternal tissues in samples isolated at the stage where specific excision of fresh endosperm was most difficult, but the relative transcript abundance was low. Indeed, comparisons with rice confirmed that many orthologs of barley transcripts enriched in modules such as green yellow at 3 DAP were expressed in rice endosperm at a similar stage, and included several putative endosperm-specific genes (Supplemental Fig. S6 and Supplemental Fig. S7C). However, the transcript abundance of these photosynthesis-related genes, which included Rubisco and genes for chlorophyll synthesis, was very low relative to other genes of interest. It was also noteworthy that the saddle brown module, which contained differentially expressed genes for starch synthesis, showed increased expression levels from 5 DAP to 7 DAP, which is consistent with the appearance of starch granules in the endosperm at 6 DAP (Supplemental Fig. S7D). In addition, the black module contained multiple genes encoding hordein storage proteins that showed increased expression from 6 DAP onwards (Supplemental Fig. S7E). These two modules provided a key reference point for the later stages of endosperm cellularization and differentiation.

To mine our transcript data more deeply, we focused on modules and module clusters that were enriched in ontological terms related to cell wall biosynthesis, remodeling, and/or degradation. These terms were enriched in several overlapping modules but clearly indicated that there were at least two stages in differential expression of genes involved in wall metabolism, namely at 3 DAP to 5 DAP, and 6 DAP to 8 DAP. Callose synthesis terms were enriched in the yellow module (Supplemental Fig. S3) and the peak of differential expression at 3 DAP to 4 DAP was consistent with the extension of the callosic anticlinal walls from the periphery of the syncytium at that stage of grain development. The overlapping red module was enriched in terms relating to embryo sac cellularization, which was again consistent with the morphological data (Fig. 1) (Wilson et al., 2006). The light cyan 1 and plum modules, which both included terms related to plant cell-wall biogenesis, cellulose synthesis, and primary cell wall synthesis, showed maximal differential expression at 6 DAP. Although the GO cell wall terms were spread across different modules, the analyses strongly suggested that there were two distinct stages of cell wall biosynthesis during early barley grain development, namely at 3 DAP to 4 DAP, and 5 DAP to 8 DAP.

There are a number of possible biological explanations for the two apparent stages in cell wall metabolism. In the early 3 DAP to 4 DAP period, anticlinal cell wall outgrowths would be expected to be associated with terms such as callose biosynthesis, and this is precisely the period in which immunocytochemical studies detected the presence of callose in the developing barley endosperm (Wilson et al., 2006). However, with the initiation of periclinal wall formation and mitosis at 5 DAP to 6 DAP, one might expect to observe a staged accumulation of transcripts in preparation for the initiation of cell division and new wall formation from 5 DAP onwards. Furthermore, Wilson and colleagues (Wilson et al., 2006, 2012) reported that important changes occur in cell wall composition at this stage, when more normal wall biosynthesis is initiated. Firstly, they detected the presence of (1,3;1,4)-β-glucans and heavily substituted arabinoxylans for the first time at about 5 DAP. These two polysaccharides continue to accumulate until grain maturity, when they constitute close to 90% of the starchy endosperm cell walls (Fincher, 1975). Wilson et al. (2006) also showed that both xyloglucans and (1,3)-β-glucans appeared transiently at 3 DAP to 4 DAP, but had disappeared or were present only in small amounts by 8 DAP to 10 DAP. Further, heteromannans also appeared transiently in the walls from 5 DAP to 6 DAP but were not detectable by 10 DAP. Thus, during this period of major transition in wall composition we would expect to see large changes in the expression levels of genes involved not only in the synthesis of new wall polysaccharides, but also in the expression of genes that mediate polysaccharide hydrolysis during turnover of the transiently appearing wall components.

The next objective was to investigate any functional or temporal linkages between the modules that were enriched in cell wall-related genes. A cell wall-specific coexpression analysis was therefore performed and revealed a network of cell wall genes that was connected by one or a few nodes (Fig. 3). There was a significant temporal element to the network, leading from the genes expressed at 3 DAP to 4 DAP to those differentially expressed later in the stages of early grain development examined. The network shown in Figure 3 illustrates the progression from genes expressed at 3 DAP to 4 DAP through to those differentially expressed at 5 DAP to 6 DAP. In the cluster of very early stage modules A, B, and C, there is a concentration of genes encoding expansins, AGPs, and a range of

polysaccharide hydrolases. Quantitative PCR confirmed that multiple genes from this group are expressed in the young endosperm (Supplemental Fig. S7). The second cluster, consisting of modules D, E, and F (Fig. 3) also contains a number of genes that encode polysaccharide hydrolases or enzymes that remodel polysaccharides during normal cell maturation. Subsequent regions of the network are characterized by an increasing abundance of glycosyl transferase (GT) genes, although genes encoding polysaccharide hydrolases are still present in this part of the network.

In most cases the presence of specific genes in this network can be rationalized with the physical and biochemical changes in cell walls that are occurring during the endosperm cellularization process. Firstly, expansins are proteins that are involved in wall loosening during cell expansion (Cosgrove, 2005), so the presence of genes encoding expansins would be expected throughout the grain development process. Secondly, AGPs have been implicated in a wide range of biological functions and have previously been shown to accumulate to high levels in the developing embryo sac tissues of many plants (Tucker and Koltunow, 2014). The AGPs and fasciclin-like AGPs of the type detected here function, inter alia, in cell expansion, protein-protein interactions, and cell type specification (Seifert and Roberts, 2007; Tan et al., 2012). All of these processes will be occurring during cellularization and it is therefore difficult to assign a more specific function to these genes at this stage. Thirdly, the relative abundance of hydrolytic enzymes in the initial gene clusters can be explained in at least two ways. At this stage of endosperm development, (1,3)-β-glucans, xyloglucans, and mannans appear transiently, as discussed above (Wilson et al., 2012), and the hydrolases are likely to function in the removal of these polysaccharides from the walls. Thus, (1,3)-β-glucan endo- and exohydrolases (GH17 and GH3, respectively), (1,4)-β-mannan endohydrolases (GH5), and members of the GH9 family, which are generally classified as cellulases but also include xyloglucan-specific (1,4)-β-glucanases, are all concentrated in the first two clusters of the network, together with a xyloglucan endotransglycosylase (XET) gene (Fig. 3). Another possible explanation for the abundance of genes encoding polysaccharide hydrolases in the network is related to emerging evidence that in most cells where cell-wall polysaccharide biosynthesis is occurring, apparently counteracting polysaccharide hydrolases occur simultaneously (Fincher, 2009). The precise role played by hydrolases in polysaccharide synthesis is unknown, but it has been suggested that the enzymes might be involved in cleavage of the nascent polysaccharide from its synthase enzyme, that they might be required to somehow trim the nascent polysaccharides, or that they might be involved in the manipulation of the molecular size distribution of the polysaccharides (Fincher, 2009). The hydrolases could also be related to a completely different function, because both (1,3)-β-glucanases and GH18 chitinases of the type seen in the G/I/K network cluster in Figure 3, are pathogenesis-related proteins that are particularly abundant in mature grain and might play a preemptive role in the protection of the developing grain against pathogen invasion (Fincher, 1989). In any case, genes encoding polysaccharide hydrolases, especially the GH17 (1,3)-β-glucanases, are found throughout the network (Fig. 3).

Consistent with the transient appearance of callose and (1,4)-β-mannans in endosperm cellularization is the presence in the early regions of the network of GT48 genes, which are likely to encode callose synthases, and GT2 *CslA* genes, some of which encode (1,4)-β-mannan synthases. In addition, the GMPP gene in the D/E/F cluster is likely to be involved in the production of the GDP-Glc nucleotide-sugar substrate for (1,4)-β-mannan biosynthesis (Fig. 3). One of the GT48 genes, *HvGSL3* (MLOC_20815), was up-regulated at 4 DAP, coincident with the appearance of callose in anticlinal endosperm walls. Despite being expressed at relatively low levels in several barley tissues, *HvGSL3* shows a unique pattern of transcript abundance compared to six other *GSL* genes (Schober et al., 2009), and may play a region-specific role within the developing endosperm. By contrast, *HvGSL5* (MLOC_64106) was up-regulated at 5 DAP and was previously identified as the most abundant GT48 gene detected during early endosperm development (Schober et al., 2009). *HvGSL5* appears more likely to play a role during the waves of cell division and cytokinesis that follow initial anticlinal wall formation.

In the second D/E/F cluster of the expression network, several genes are found that are potentially involved in pectic polysaccharide synthesis (family GT8) and degradation (PME, pectin acetylesterase, and polygalacturonase; Fig. 3). While pectin has generally been assumed to be absent from endosperm cell walls of barley endosperm and most other cereals (Stone and Fincher, 2004), it has been reported to be found in the walls of rice endosperm (Ishii et al., 1989) and, more recently, in the starchy endosperm walls of wheat, albeit at low levels (Chateigner-Boutin et al., 2014). The differential expression of these genes during cellularization of the barley endosperm suggests that pectin is deposited transiently in the walls or in cell plates during mitotic cell division.

Low levels of cellulose are also detected in endosperm cell walls at 3 DAP and 4 DAP (Wilson et al., 2012), but cellulose is difficult to distinguish from other (1,4)-β-glucans at later stages. In the walls of mature barley grains, cellulose accounts for just 3% to 4% by weight of the walls (Fincher, 1975). Two cellulose synthase (*CesA*) genes were identified in the module analysis, including *HvCesA6* (plum 2), which is often transcribed during primary cell wall biosynthesis, and *HvCesA3* (red). Both genes showed a peak in transcript abundance at 5 DAP, but coexpression values that were too low to be included in the top 5% of correlations for spatial molecular network analysis.

Turning now to the two most abundant polysaccharides in the endosperm cell walls of mature barley

grain, namely the (1,3;1,4)-β-glucans and the arabinoxylans, it is evident from our analyses that genes that have been implicated in their biosynthesis are not heavily represented in the differentially expressed gene lists. The *HvCslF9* gene, which is believed to encode a (1,3;1,4)-β-glucan synthase and is known to be expressed transiently at the late stages of cellularization, is represented at this time and falls into the dark gray module (Supplemental Fig. S2). The *CslF6* gene is transcribed to high levels during all endosperm stages analyzed here, consistent with previous studies of the (1,3;1,4)-β-glucan synthase family (Burton et al., 2008), but is not represented in any of the modules. This is almost certainly because the levels of the *CslF6* transcript gradually increase only after cellularization is complete and thereafter remain high throughout grain development. Thus, the gene would not be included because of the high threshold imposed here on the definition of a differentially expressed gene and the fact that only the top 5% of correlations are shown in Figure 3. The *CslH1* gene, which is also known to encode a (1,3;1,4)-β-glucan synthase (Doblin et al., 2009), can be seen in the lower region of the gene network (Fig. 3).

For the same reason, it might be expected that genes that mediate arabinoxylan biosynthesis would be underrepresented in the analysis. Nevertheless, several GT47, GT31, and GT8 genes, which have been implicated in arabinosyl and xylosyl transferase activity (Mitchell et al., 2007; Zeng et al., 2010; Anders et al., 2012; Chiniquy et al., 2012; Bromley et al., 2013; Urbanowicz et al., 2014), are detected in the network analysis (Fig. 3), as are GH10 genes that might encode endo-xylanases. In the first cluster of genes in the network a GH51 putative arabinoxylan arabinofuranohydrolase is detected and the proposed role of this enzyme in the removal of arabinosyl substituents from the (1,4)-β-xylan backbone of the arabinoxylan (Lee and Schiefelbein, 2001) is consistent with the observation of Wilson et al. (2012) that newly synthesized arabinoxylans in barley are heavily substituted with arabinosyl residues, but that these substituents are subsequently removed.

Our final analysis was aimed at identifying TF genes that might be involved in the regulation of cell wall synthesis, remodeling, or degradation during the multiple changes that occur in wall biology during endosperm cellularization. One question of special relevance in the developing endosperm, where cell walls are not subject to secondary thickening, was whether TF repressors of secondary wall synthesis were expressed; many TFs that control secondary wall biosynthesis belong to the NAC or MYB groups (Zhou et al., 2009; Yao et al., 2012; Valdivia et al., 2013; Zhong and Ye, 2015; Zhou et al., 2014; Taylor-Teeples et al., 2015). Four putative MYB TFs and one NAC TF showed varied differential patterns across the 6 d time course. However, secondary cell wall genes such as Group 2 *CesA*s or lignin biosynthetic genes were not detected in any of the differential cell wall lists, suggesting that the NAC and MYB TF genes identified here might be linked to roles in primary rather than secondary cell wall biogenesis.

A second coexpression network was generated using a set of TFs identified in the barley genome reference assembly (Mayer et al., 2012) in combination with the differential subset of cell wall genes that were examined here (Supplemental Table S6, Fig. 5). Although this coexpression network may not necessarily reveal direct interactions between TF and cell wall genes, it highlights TFs that are associated with specific stages of endosperm development and may be considered candidate genes for future studies on TFs that influence developmental events of interest. The shape of the cell wall-specific network changed with the addition of the TF genes, but the temporal connections between clusters were still evident, and many of the associations between cell wall genes seen in Figure 3 remained. In some cases the TF genes created linkages between previously unconnected groups. Four clusters were identified that were of particular interest.

In several cases, genes from the clusters were found to share partially overlapping expression patterns not only during early endosperm development, but also in RNAseq transcript data previously generated in the barley genome project from various barley tissue extracts during general growth and development (Fig. 5; Mayer et al., 2012). Several of the TF/cell wall gene combinations were also coexpressed across development in other grass species such as *Brachypodium* (Supplemental Fig. S5, G and H). A cluster of MADS genes is seen in the early stages of cellularization (A/C cluster in Fig. 5), and this is linked with AGP, expansin, GT48 callose synthase, and GH16 XET genes. A separate cluster that includes MADS, TFID, HD-ZIP, and BTF3-like TFs (cluster B in Fig. 5) is associated with many genes, including those encoding expansins, AGPs, mannan synthases (*CslA*), GH5 mannanases, and (1,3)-β-glucanases (Fig. 5). Transcript accumulation of the MADS TF and GH9 genes (Fig. 5E) was almost identical across development, while HD-ZIP TF and GH17 (Fig. 5F) and MADS TF and AGP/fasciclin-like AGP (Fig. 5G) displayed similar expression patterns and peaked during inflorescence or early caryopsis development. Transcription of the putative ethylene response factor and GMPP (Fig. 5H) was also highly correlated during leaf, root, inflorescence, caryopsis, and embryo development. These data indicate that some of the associations between cell wall-related genes and TFs in the early endosperm data may relate to more general associations across development. These examples illustrate the potential of the dataset and the network analyses not only to identify TFs that are associated with key stages of cell wall metabolism in a grain-specific context, but also to provide more general clues as to which TFs might influence or control specific aspects of wall metabolism in other tissues of the plant.

In summary, we have collected large transcript abundance datasets from developing barley grain that span the period of syncytium cellularization. From 3 DAP to 8 DAP, several dramatic changes occur in the cell biology of the endosperm, including the growth of cell walls around free nuclei to cellularize the endosperm,

a switch to mitotic cell division, the expansion of the endosperm, the differentiation of the aleurone layer, and the deposition of starch and storage protein. To demonstrate the potential value of the analyses, the data have been mined for information on the transcription patterns of genes that are involved in cell wall biosynthesis and these data are superimposable on cellular and biochemical events that are shown to occur at the same time. Spatial and temporal coexpression networks have been defined and TFs that putatively link the expression patterns of groups of genes have been identified. Finally, the data have been made publicly available at https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-2728/ and through a dedicated web interface (https://ics.hutton.ac.uk/barseed/). We see the datasets as a global resource through which coexpression and differential expression patterns for any genes of interest can be interrogated and as a method for identifying TFs that might control individual genes, biochemical pathways, or entire cellular processes.

MATERIALS AND METHODS

Plant Material and RNA Isolation

Endosperm was isolated from developing *H. vulgare* cv Sloop grain essentially as described by Burton et al. (2008). Spikes were hand-pollinated at anthesis and developing grains were collected from 0 DAP to 8 DAP. For material, 3–6 DAP, after removal of the very base of the caryopsis to excise the young embryo, the liquid endosperm was aspirated out of the developing caryopses following the application of gentle pressure and immediately frozen in liquid nitrogen. For 0, 1, and 2 DAP samples, whole ovaries were harvested, because of the difficulty in extracting the small embryo sacs and syncytial endosperm, but these preparations were not used in the central analyses of the work. After 6 DAP, cellularized endosperm was collected by removing the maternal pericarp tissue, which separated cleanly leaving no residual cells, and excising the basal region containing the embryo. Endosperm from approximately 50 caryopses from at least three spikes was collected and combined for RNA extraction for each biological replicate.

Total RNA was prepared from tissue homogenates using a commercially prepared phenol-guanidine reagent (Trizol; Life Technologies, Carlsbad, CA) according to the manufacturer's guidelines and pretreated with the DNA-free kit (Ambion, Austin, TX) prior to its use as the template for cDNA synthesis using Superscript III reverse transcriptase (Life Technologies), as previously described in Burton et al., 2011.

Microarray Experiments

A custom microarray SCRI_Hv35_44k_v2 (Agilent design 023660; ArrayExpress accession no. A-MEXP-2404; Agilent Technologies, Santa Clara, CA) representing 42,908 barley sequences was generated. Barley sequences for this design were selected from a total of 50,938 unigenes from HarvEST assembly 35 (http://www.harvest-web.org/) representing approximately 450,000 expressed sequence tags. Selection criteria were based upon the ability to define orientation derived from: (i) homology to members of the nonredundant protein database (NCBI nr); (ii) homology to expressed sequence tags known to originate from directional cDNA libraries; and (iii) presence of a significant polyA tract. In addition, 552 selected publicly available cell wall-related genes were added to the design. The microarray was designed with one 60-mer probe per selected unigene in 4 × 44 K format using default parameters in the web-based Agilent eArray software (https://earray.chem.agilent.com/earray/) and includes recommended QC control probes.

The experiment was run using recommended protocols for processing Agilent microarrays. The quality of the total RNA preparation was first checked using a Bioanalyzer, and labeled using the One-Color Low Input Labeling Kit (Agilent). The experimental design was as described (ArrayExpress accession no. E-MTAB-2728) with one sample per array. Hybridization, washing, and scanning were performed as recommended. Probe data were extracted from individual images using Feature Extraction software (v. 10.7; Agilent), QC checked and uploaded to Genespring (v. 7.3; Agilent). Data were globally normalized in Genespring v. 7.3 using default parameters for one-color data: (i) each probe signal was set to a minimum of 5; (ii) values for each microarray were normalized to the median of signal within the microarray; and (iii) values for each probe were normalized the median of its signal across all microarrays. Probes with at least two present or marginal flags within three biological replicates in at least one stage were selected, which reduced the total number of probes from 42,908 to 27,508. The processed data were used for statistical analyses.

Identification of Preferentially and Differentially Expressed Transcripts

Present or marginal flags for each probe were used to indicate expression in a tissue. Within the three biological replicates at a certain stage, a probe with at least two present or marginal calls was considered as expressed at that stage. The probes that are expressed in one stage, but not expressed in any other stages in the experiment, are deemed as preferentially expressed transcripts. For the coexpression network analysis and as a basic gene list for the pairwise t-tests, 5925 probes showing significantly different expression between any two time points in the six stages of endosperm development, using an ANOVA model and an false discovery rate threshold of 1×10^{-4}, were selected. A Welch's T-test was used to detect significant differentially expressed genes between adjacent time points (e.g. 4 DAP vs. 3 DAP, 5 DAP vs. 4 DAP, etc.) using a p-value threshold of 0.01 on the ANOVA filtered list mentioned above.

Gene Ontology Enrichment Analysis

Gene ontology (GO) enrichment of the gene lists was carried out using the TopGO package (v. 2.12) in R (v. 3.1.0). By taking advantage of the GO annotations for barley Morex locuses (MLOCs) in a previous study (Baker et al., 2014), we aligned microarray probe sequences against barley MLOC transcripts using BLAST (Altschul et al., 1990). Barley MLOC hits with an e-value cutoff of 1e-20 for each probe were selected. Finally, probes were annotated with GO terms from the associated barley MLOCs. In total, 78,449 pairs of probe and GO term relationships have been identified for 16,911 probes, which could be summarized into 3917 GO terms. All probes on the microarray with GO annotations were used as background and genes with GO annotations from generated lists were investigated for enriched GO terms. Only GO terms with at least five annotated genes were considered. A Fisher statistics and weight algorithm was used to calculate the p-values of significantly enriched GO terms. A p-value of 0.05 was used as a cutoff to select GO terms.

Coexpression Network Construction

A gene coexpression network was constructed using the R weighted gene coexpression analysis (WGCNA v. 1.27-1) library in R (v 3.1.0; Gentleman et al., 2003). A gene expression adjacency matrix was constructed using the soft thresholded Pearson correlation (power = 30) between all pairs of probe expression data. Such soft thresholding yields a scale-free network (scale-free topology model fit $R2 = 0.8$) with an average connectivity of 380 (median 257). A topological overlap map was calculated by comparing the similarities of connectivity of each pair of probes, in which the strength of connectivity is defined in the adjacency matrix. A dynamic hybrid tree-cut algorithm (R package dynamicTreeCut, v. 1.60) was used for module detection. The minimum size of the module was set to 30 to avoid small modules and DeepSplit was set to 3 to produce fine clusters. The modules with high similarity of eigengenes (dissimilarity <0.025) have been merged to avoid over-splitting clusters. Each detected module is represented by color-coding, with 48 modules detected. The gray module, reserved for probes that did not fit any other module, was removed from further analysis.

Network Visualization

For the cell wall-related subnetwork, a Spearman correlation matrix was calculated based on the 89 probes related to cell walls. The probes whose correlations with other probes ranked in the top 5% were selected for the visualization of coexpressed gene network in Cytoscape (v. 3.2.1; Shannon et al., 2003).

The same approach was used for the transaction-factor-associated cell wall subnetwork. A total of 57 putative transaction-factor-encoding genes, together with the 89 cell wall genes, were used to create a network map and the top 1% of Spearman correlations was investigated manually in detail.

Supplemental Data

The following supplemental materials are available.

Supplemental Fig. S1. Regions of barley grain harvested for microarray analysis.

Supplemental Fig. S2. PCA of barley endosperm microarray data.

Supplemental Fig. S3. Modules and enriched GO categories.

Supplemental Fig. S4. Spatial representation of a 5–8 DAP cell-wall network.

Supplemental Fig. S5. Transcript profiles of barley MLOCs and *Brachypodium* orthologs.

Supplemental Fig. S6. Heatmap showing normalized expression levels of 550 rice genes in ovary, caryopsis, embryo, endosperm, and leaf tissues.

Supplemental Fig. S7. Transcript profiles of putative endosperm-specific, photosynthesis-related, hordein, and starch synthase genes.

Supplemental Fig. S8. Transcript profiles of the *HvEX1* and *HvNUC1* nucellus marker genes.

Supplemental Table S1. All differentially expressed probes.

Supplemental Table S2. Differential probe GO analysis.

Supplemental Table S3. Preferentially expressed probes.

Supplemental Table S4. Preferential probe GO analysis.

Supplemental Table S5. Module GO analysis.

Supplemental Table S6. Normalized microarray data including 0 DAP, 1 DAP, and 2 DAP samples.

Supplemental Table S7. Differentially expressed cell wall-related genes.

Supplemental Table S8. Differentially expressed TF-encoding genes.

Supplemental Table S9. Quality assessment of differential barley probes compared to published datasets.

ACKNOWLEDGMENTS

The authors acknowledge Sebastian Raubach for helping with the expression browser, Katie Baker for providing the GO annotations for barley MLOCs, and Jessica Gilson for expert technical assistance.

Received November 2, 2015; accepted January 9, 2016; published January 11, 2016.

LITERATURE CITED

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol **215:** 403–410

Anders N, Wilkinson MD, Lovegrove A, Freeman J, Tryfona T, Pellny TK, Weimar T, Mortimer JC, Stott K, Baker JM, Defoin-Platel M, Shewry PR, et al (2012) Glycosyl transferases in family 61 mediate arabinofuranosyl transfer onto xylan in grasses. Proc Natl Acad Sci USA **109:** 989–993

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, et al; The Gene Ontology Consortium (2000) Gene ontology: tool for the unification of biology. Nat Genet **25:** 25–29

Bacic A, Stone B (1981) Chemistry and organization of aleurone cell wall components from wheat and barley. Funct Plant Biol **8:** 475–495

Baker K, Bayer M, Cook N, Dreißig S, Dhillon T, Russell J, Hedley PE, Morris J, Ramsay L, Colas I, Waugh R, Steffenson B, et al (2014) The low-recombining pericentromeric region of barley restricts gene diversity and evolution but not gene expression. Plant J **79:** 981–992

Becraft PW, Asuncion-Crabb Y (2000) Positional cues specify and maintain aleurone cell fate in maize endosperm development. Development **127:** 4039–4048

Bromley JR, Busse-Wicher M, Tryfona T, Mortimer JC, Zhang Z, Brown DM, Dupree P (2013) GUX1 and GUX2 glucuronyltransferases decorate distinct domains of glucuronoxylan with different substitution patterns. Plant J **74:** 423–434

Brown RC, Lemmon BE, Olsen O-A (1994) Endosperm development in barley: microtubule involvement in the morphogenetic pathway. Plant Cell **6:** 1241–1252

Burton RA, Collins HM, Kibble NA, Smith JA, Shirley NJ, Jobling SA, Henderson M, Singh RR, Pettolino F, Wilson SM, Bird AR, Topping DL, et al (2011) Over-expression of specific HvCslF cellulose synthase-like genes in transgenic barley increases the levels of cell wall (1,3;1,4)-β-d-glucans and alters their fine structure. Plant Biotechnol J **9:** 117–135

Burton RA, Jobling SA, Harvey AJ, Shirley NJ, Mather DE, Bacic A, Fincher GB (2008) The genetics and transcriptional profiles of the cellulose synthase-like HvCslF gene family in barley. Plant Physiol **146:** 1821–1833

Chateigner-Boutin A-L, Bouchet B, Alvarado C, Bakan B, Guillon F (2014) The wheat grain contains pectic domains exhibiting specific spatial and development-associated distribution. PLoS One **9:** e89620

Chiniquy D, Sharma V, Schultink A, Baidoo EE, Rautengarten C, Cheng K, Carroll A, Ulvskov P, Harholt J, Keasling JD, Pauly M, Scheller HV, et al (2012) XAX1 from glycosyltransferase family 61 mediates xylosyltransfer to rice xylan. Proc Natl Acad Sci USA **109:** 17117–17122

Cosgrove DJ (2005) Growth of the plant cell wall. Nat Rev Mol Cell Biol **6:** 850–861

Doan DNP, Linnestad C, Olsen OA (1996) Isolation of molecular markers from the barley endosperm coenocyte and the surrounding nucellus cell layers. Plant Mol Biol **31:** 877–886

Doblin MS, Pettolino FA, Wilson SM, Campbell R, Burton RA, Fincher GB, Newbigin E, Bacic A (2009) A barley cellulose synthase-like CSLH gene mediates (1,3;1,4)-beta-D-glucan synthesis in transgenic Arabidopsis. Proc Natl Acad Sci USA **106:** 5996–6001

Fincher G (1975) Morphology and chemical composition of barley endosperm cell walls. J Inst Brew **81:** 116–122

Fincher GB (1989) Molecular and cellular biology associated with endosperm mobilization in germinating cereal grains. Annu Rev Plant Biol **40:** 305–346

Fincher GB (2009) Revolutionary times in our understanding of cell wall biosynthesis and remodeling in the grasses. Plant Physiol **149:** 27–37

Fincher GB, Stone B (2004) Chemistry of non-starch polysaccharides from cereal grains. *In* Y. Pomeranz, ed, Advances in Cereal Science and Technology, **Vol III**. American Association of Cereal Chemists, St. Paul, MN

Gentleman R, Ihaka R, Leisch F (2003) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria

Girke T, Lauricha J, Tran H, Keegstra K, Raikhel N (2004) The Cell Wall Navigator database. A systems-based approach to organism-unrestricted mining of protein families involved in cell wall metabolism. Plant Physiol **136:** 3003–3008

Ishii T, Thomas J, Darvill A, Albersheim P (1989) Structure of plant cell walls: XXVI. The walls of suspension-cultured sycamore cells contain a family of rhamnogalacturonan-I-like pectic polysaccharides. Plant Physiol **89:** 421–428

Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics **9:** 559

Lee MM, Schiefelbein J (2001) Developmentally distinct MYB genes encode functionally equivalent proteins in Arabidopsis. Development **128:** 1539–1546

Linnestad C, Doan DNP, Brown RC, Lemmon BE, Meyer DJ, Jung R, Olsen OA (1998) Nucellain, a barley homolog of the dicot vacuolar-processing protease, is localized in nucellar cell walls. Plant Physiol **118:** 1169–1180

Mayer KF, Waugh R, Brown JW, Schulman A, Langridge P, Platzer M, Fincher GB, Muehlbauer GJ, Sato K, Close TJ, Wise RP, Stein N; International Barley Genome Sequencing Consortium (2012) A physical, genetic and functional sequence assembly of the barley genome. Nature **491:** 711–716

Mitchell RA, Dupree P, Shewry PR (2007) A novel bioinformatics approach identifies candidate genes for the synthesis and feruloylation of arabinoxylan. Plant Physiol **144:** 43–53

Nie DM, Ouyang YD, Wang X, Zhou W, Hu CG, Yao J (2013) Genome-wide analysis of endosperm-specific genes in rice. Gene **530:** 236–247

Olsen OA (2004) Nuclear endosperm development in cereals and Arabidopsis thaliana. Plant Cell **16**(Suppl): S214–S227

Paulson I (1969) Embryogeny and caryopsis development of Sorghum bicolor (L.) Moench. Crop Sci **9:** 97–102

Pellny TK, Lovegrove A, Freeman J, Tosi P, Love CG, Knox JP, Shewry PR, Mitchell RA (2012) Cell walls of developing wheat starchy endosperm: comparison of composition and RNA-Seq transcriptome. Plant Physiol **158:** 612–627

Pfeifer M, Kugler KG, Sandve SR, Zhan B, Rudi H, Hvidsten TR; International Wheat Genome Sequencing Consortium, Mayer KF, Olsen OA (2014) Genome interplay in the grain transcriptome of hexaploid bread wheat. Science **345:** 1250091

Sato Y, Takehisa H, Kamatsuki K, Minami H, Namiki N, Ikawa H, Ohyanagi H, Sugimoto K, Antonio BA, Nagamura Y (2013) RiceXPro version 3.0: expanding the informatics resource for rice transcriptome. Nucleic Acids Res **41:** D1206–D1213

Schober MS, Burton RA, Shirley NJ, Jacobs AK, Fincher GB (2009) Analysis of the (1,3)-β-D-glucan synthase gene family of barley. Phytochemistry **70:** 713–720

Seifert GJ, Roberts K (2007) The biology of arabinogalactan proteins. Annu Rev Plant Biol **58:** 137–161

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res **13:** 2498–2504

Sreenivasulu N, Borisjuk L, Junker BH, Mock HP, Rolletschek H, Seiffert U, Weschke W, Wobus U (2010) Barley grain development toward an integrative view. Int Rev Cell Mol Biol **281:** 49–89

Sreenivasulu N, Usadel B, Winter A, Radchuk V, Scholz U, Stein N, Weschke W, Strickert M, Close TJ, Stitt M, Graner A, Wobus U (2008) Barley grain maturation and germination: metabolic pathway and regulatory network commonalities and differences highlighted by new MapMan/PageMan profiling tools. Plant Physiol **146:** 1738–1758

Sturaro M, Linnestad C, Kleinhofs A, Olsen OA, Doan DNP (1998) Characterization of a cDNA encoding a putative extensin from developing barley grains (*Hordeum vulgare* L.). J Exp Bot **49:** 1935–1944

Tan L, Showalter AM, Egelund J, Hernandez-Sanchez A, Doblin MS, Bacic A (2012) Arabinogalactan-proteins and the research challenges for these enigmatic plant cell surface proteoglycans. Front Plant Sci **3:** 140

Taylor-Teeples M, Lin L, de Lucas M, Turco G, Toal TW, Gaudinier A, Young NF, Trabucco GM, Veling MT, Lamothe R, Handakumbura PP, Xiong G, et al (2015) An Arabidopsis gene regulatory network for secondary cell wall synthesis. Nature **517:** 571–575

Thiel J, Hollmann J, Rutten T, Weber H, Scholz U, Weschke W (2012) 454 Transcriptome sequencing suggests a role for two-component signalling in cellularization and differentiation of barley endosperm transfer cells. PLoS One **7:** e41867

Trafford K, Haleux P, Henderson M, Parker M, Shirley NJ, Tucker MR, Fincher GB, Burton RA (2013) Grain development in Brachypodium and other grasses: possible interactions between cell expansion, starch deposition, and cell-wall synthesis. J Exp Bot **64:** 5033–5047

Tucker MR, Koltunow AM (2014) Traffic monitors at the cell periphery: the role of cell walls during early female reproductive cell differentiation in plants. Curr Opin Plant Biol **17:** 137–145

Urbanowicz BR, Peña MJ, Moniz HA, Moremen KW, York WS (2014) Two Arabidopsis proteins synthesize acetylated xylan in vitro. Plant J **80:** 197–206

Valdivia ER, Herrera MT, Gianzo C, Fidalgo J, Revilla G, Zarra I, Sampedro J (2013) Regulation of secondary wall synthesis and cell death by NAC transcription factors in the monocot Brachypodium distachyon. J Exp Bot **64:** 1333–1343

Wilson SM, Burton RA, Collins HM, Doblin MS, Pettolino FA, Shirley N, Fincher GB, Bacic A (2012) Pattern of deposition of cell wall polysaccharides and transcript abundance of related cell wall synthesis genes during differentiation in barley endosperm. Plant Physiol **159:** 655–670

Wilson SM, Burton RA, Doblin MS, Stone BA, Newbigin EJ, Fincher GB, Bacic A (2006) Temporal and spatial appearance of wall polysaccharides during cellularization of barley (Hordeum vulgare) endosperm. Planta **224:** 655–667

Yao D, Wei Q, Xu W, Syrenne RD, Yuan JS, Su Z (2012) Comparative genomic analysis of NAC transcriptional factors to dissect the regulatory mechanisms for cell wall biosynthesis. BMC Bioinformatics **13** (Suppl 15): S10

Young TE, Gallie DR (2000) Programmed cell death during endosperm development. Plant Mol Biol **44:** 283–301

Zeng W, Jiang N, Nadella R, Killen TL, Nadella V, Faik A (2010) A glucurono(arabino)xylan synthase complex from wheat contains members of the GT43, GT47, and GT75 families and functions cooperatively. Plant Physiol **154:** 78–97

Zhong R, Ye Z-H (2015) Secondary cell walls: biosynthesis, patterned deposition and transcriptional regulation. Plant Cell Physiol **56:** 195–214

Zhou J, Lee C, Zhong R, Ye Z-H (2009) MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in Arabidopsis. Plant Cell **21:** 248–266

Zhou J, Zhong R, Ye Z-H (2014) Arabidopsis NAC domain proteins, VND1 to VND5, are transcriptional regulators of secondary wall biosynthesis in vessels. PLoS One **9:** e105726

Supplementary Figures

Figure S1

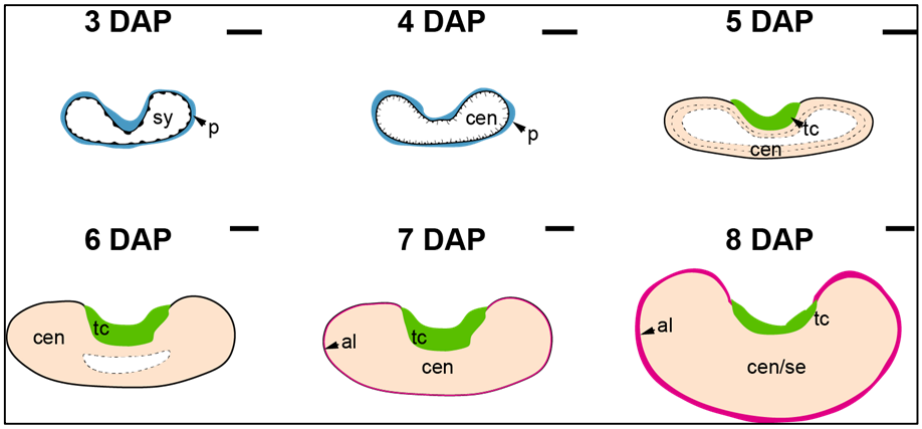

Figure S1: Diagram showing regions of barley grain manually harvested for microarray analysis (compare with Figure 1). Care was taken to avoid pericarp contamination in all stages, particularly in the 3 DAP sample. Abbreviations are as follows: p, pericarp, sy, syncytium, cen, cellular endosperm, tc, transfer cells, al, aleurone, se, starchy endosperm, DAP, days after pollination. Bar = 250 µm unless otherwise indicated.

Figure S2

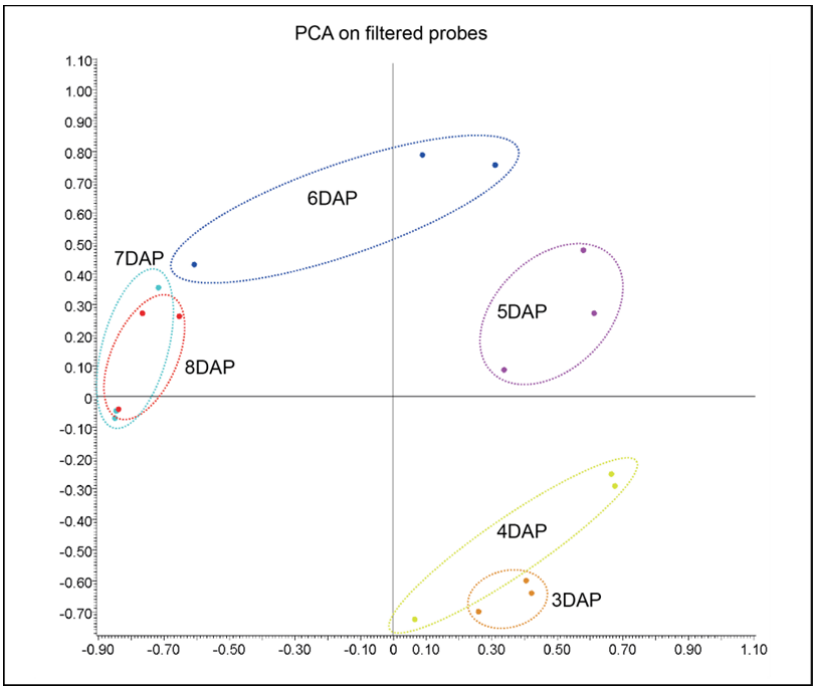

Figure S2: A primary analysis of the mRNA expression profiles for endosperm samples using principle component analysis (PCA).

Figure S3

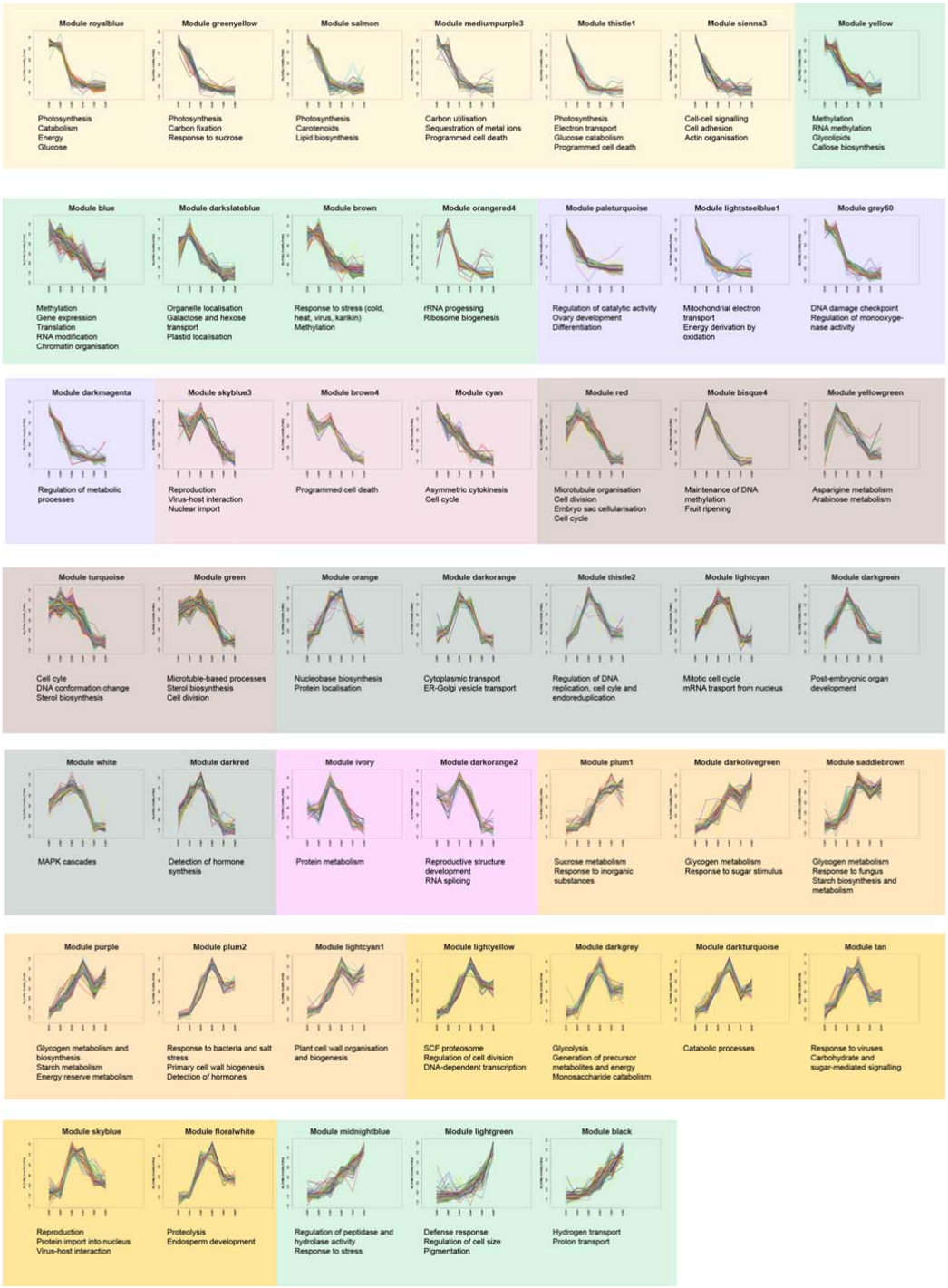

Figure S3: Enriched functional categories (based on Gene Ontology analysis) within 47 expression modules identified during early endosperm development in barley. The modules are arranged in the order shown in the dendogram (Fig. 2B). Different shading highlights the 10 main clusters of modules identified in the dendogram.

Figure S4

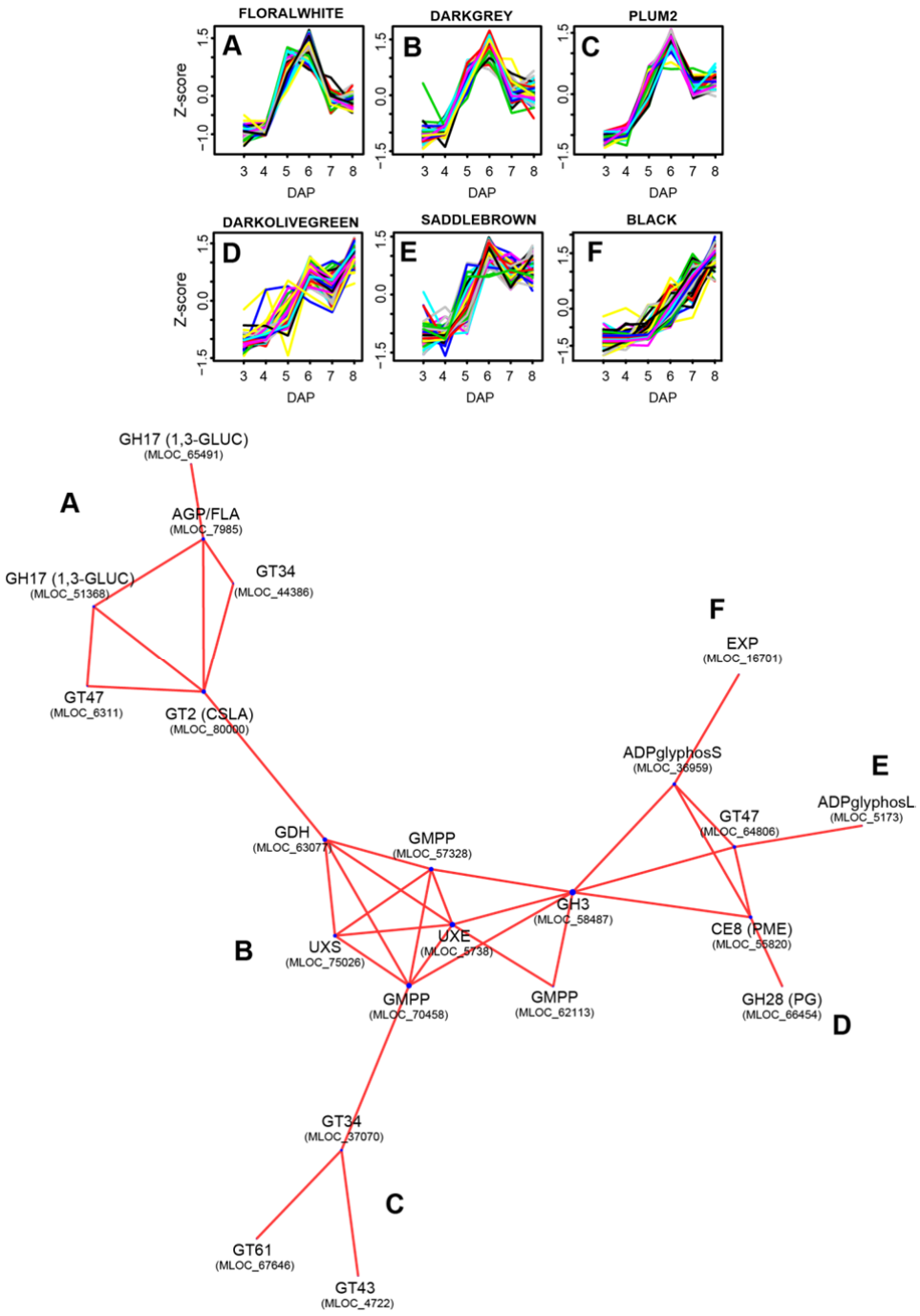

Figure S4: A spatial representation of co-expressed cell wall-related genes that are differentially expressed during 5-8 DAP of early endosperm development. Panels **A-L** show different modules that are represented by genes in the network analysis. The relative position of these modules is indicated on the map.

Figure S5

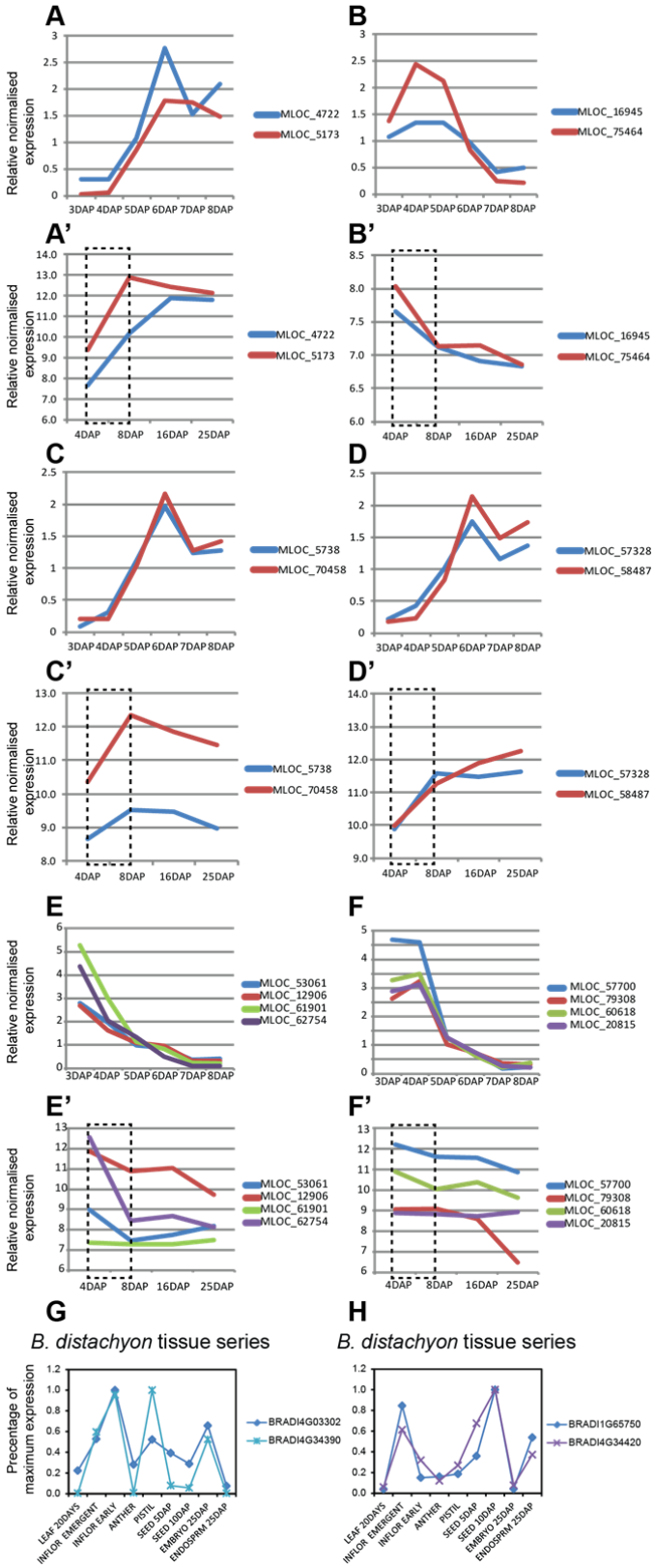

Figure S5: Transcript profiles of barley MLOCs during endosperm development (**A-F'**) and *Brachypodium* genes during general development (**G-H**). Panels **A-F** show expression profiles of MLOCs from modules identified in this study by 44K microarray analysis. Panels **A'-F'** show corresponding MLOCs extracted from 22K microarray data (Sreenivasulu et al., 2008). Although there is only a partial overlap in staging (see black dashed box) genes tended to show similar behaviour. The 44K data provides greater resolution of transcript dynamics in several cases. **G-H** show examples of two transcription factor/cell wall gene pairs that are co-expressed during general growth and development in *Brachypodium*. This is similar to the barley orthologues, although the exact tissues differ between experiments (see Fig. 5F and H)

Figure S6

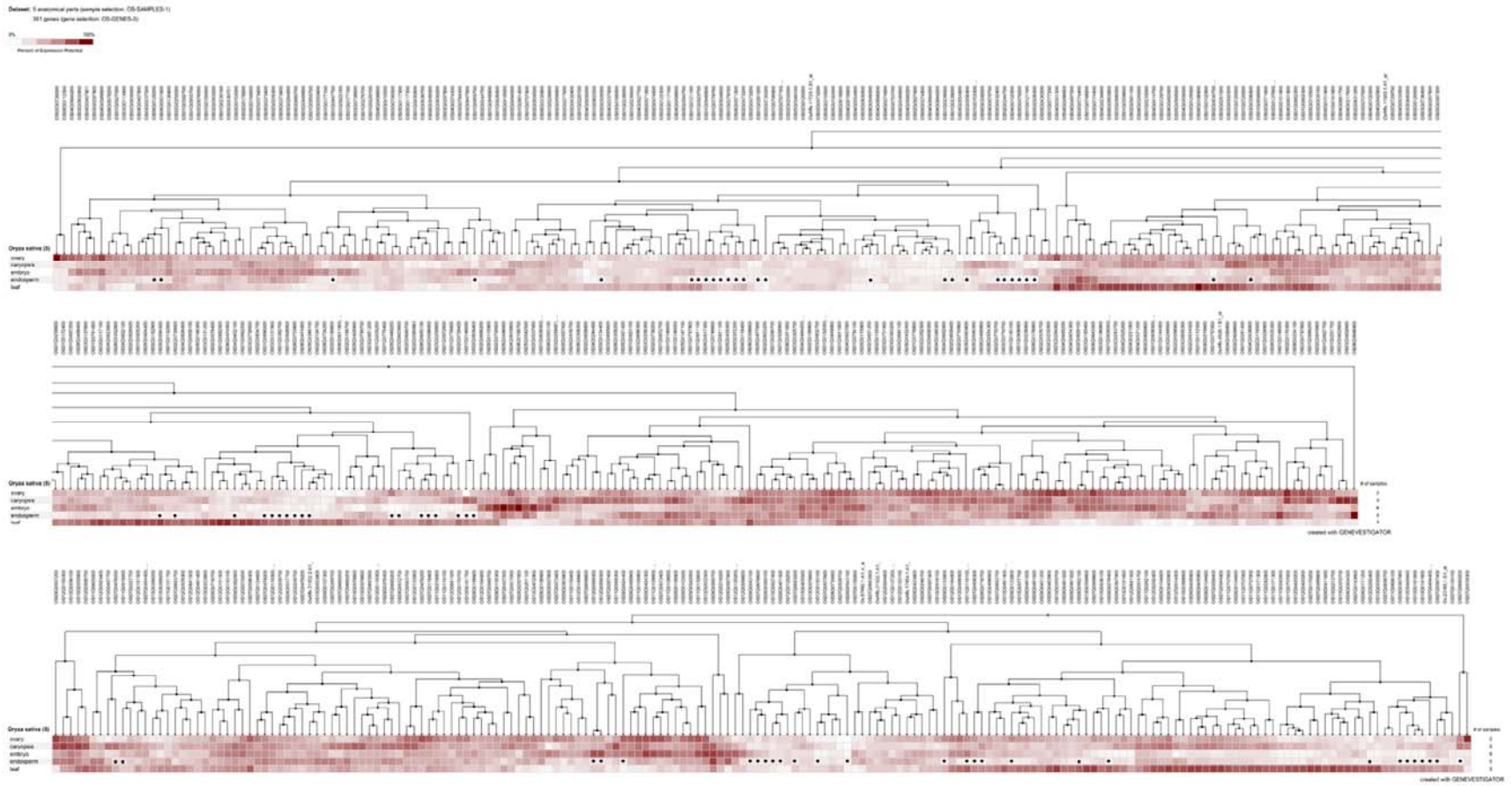

Figure S6: Heatmap showing normalised expression levels of 550 rice genes in ovary, caryopsis, embryo, endosperm and leaf tissues. The genes encode putative rice orthologues of barley genes that are differentially expressed in 3-8 DAP early endosperm time course and show peak transcript abundance at 3DAP, but are not detected in LAM data from Thiel et al., 2012. Data were extracted from Genevestigator (https://genevestigator.com/gv/doc/intro_plant.jsp). Black dots indicate samples where transcript was not detected in the endosperm. Hierarchical clustering was used to group the genes in two separate experiments (1: Rows 1&2 and 2: Row 3).

Figure S7

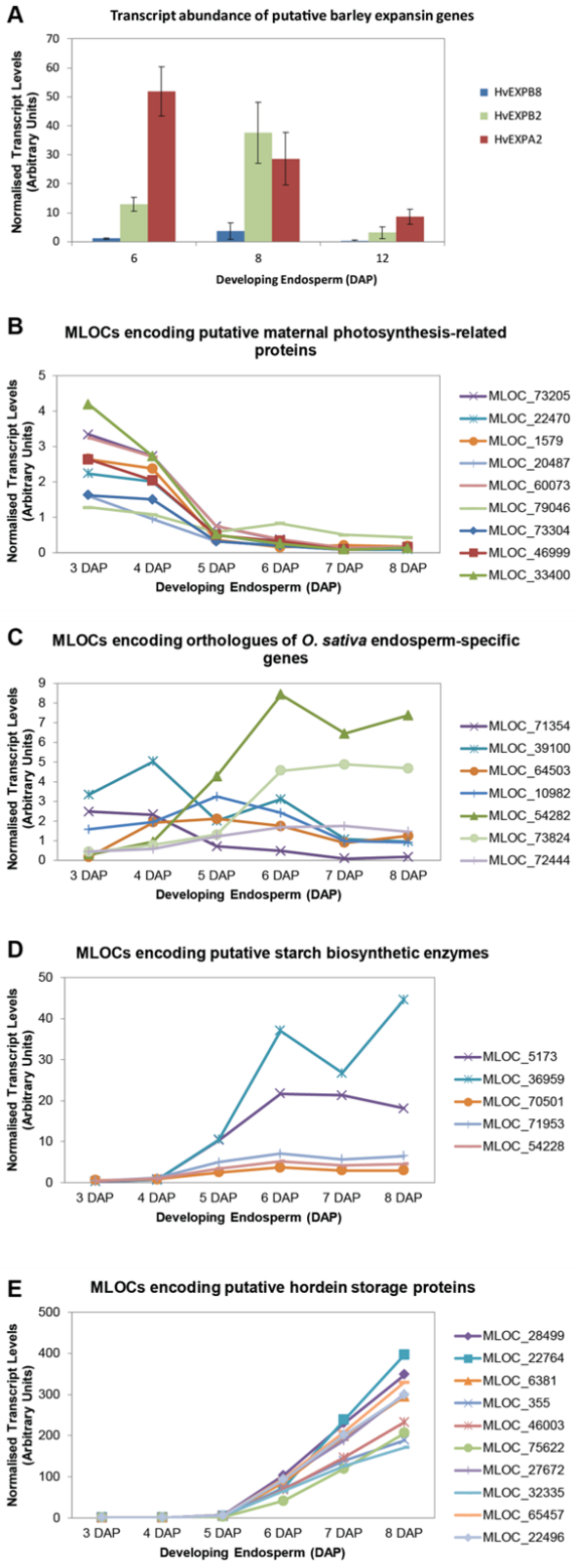

Figure S7: Transcript profiles of putative endosperm-specific, photosynthesis-related, hordein and starch synthase genes. (**A**) Quantitative PCR analysis of three putative expansin genes that are differentially expressed during barley endosperm development. (**B**) Microarray expression profiles of putative photosynthesis-related genes in the endosperm samples. (**C**) Microarray expression profiles of seven barley genes that are differentially expressed during endosperm development and whose orthologues in rice are predicted to be endosperm-specific. (**D**) Microarray expression profiles of barley genes encoding hordein storage proteins. (**E**) Microarray expression profiles of barley genes encoding starch biosynthetic enzymes.

Figure S8

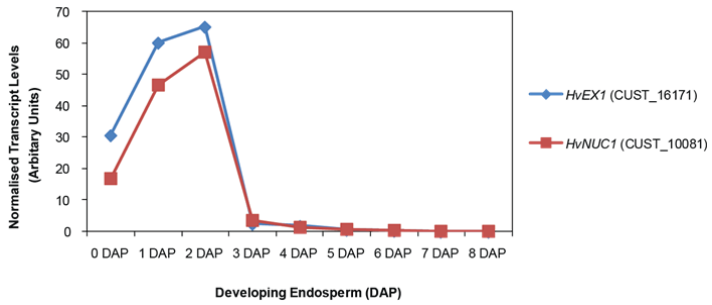

Figure S8: Transcript profiles of the *HvEX1* and *HvNUC1* nucellus marker genes identified in previous studies (Doan et al., 1996; Sturaro et al., 1998; Linnestad et al., 2015). *HvNUC1* hybridizes to transcripts in the nucellus before fertilization and in autolyzing nucellus cells after fertilization. At later stages, after the disappearance of nucellus, *HvNUC1* transcripts are present in the nucellar epidermis and in the lateral cells of the nucellar projection. *HvEX1* is expressed in grains at 8 DAP in the nucellar epidermis (ne), the nucellar projection and the ventral vascular strand.